AD-A192 434

(1)

AAMRL-TR-87-072

AD-A192 434

DTIC FILE COPY

UTILITY OF SHARED VERSUS ISOLATED WORK SETTING FOR DYNAMIC TEAM DECISION-MAKING (U)

DENISE L. WILSON
MICHAEL D. McNEESE
ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY

CLIFFORD E. BROWN
SYSTEMS RESEARCH LABORATORIES, INC.

A. RODNEY WELLENS
UNIVERSITY OF MIAMI

DECEMBER 1987

FINAL REPORT FOR PERIOD APRIL 1986 - AUGUST 1987

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
HUMAN SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

88 3 14 014

NOTICES

TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-87-072

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AAMRL-TR-87-072	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Armstrong Aerospace Medical Research Laboratory 6b. OFFICE SYMBOL (If applicable) HED	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Wright-Patterson Air Force Base OH 45433-6573	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
62202F	7184	27	03

11. TITLE (Include Security Classification)

Utility of Shared Versus Isolated Work Setting for Dynamic Team Decision-Making (U)

12. PERSONAL AUTHOR(S)

Wilson, Denise L., McNeese, Michael D., Brown, Clifford E.*, and Wellens, A. Rodney**

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Apr 86 TO Aug 87	1987 December	116

16. SUPPLEMENTARY NOTATION

*Systems Research Laboratories, Inc.

**University of Miami

Keywords:

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
23	14	
02	05	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Group Displays, Man-Machine Interface, Command and Control, Group Decision-Making, Command Centers, Command Control and Communications.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The utility of shared versus isolated operator/display setting was examined in the context of a complex, dynamic, team decision-making task. Both alphanumeric and graphic display formats were utilized as well as moderate and fast information presentation rates. Performance scores were significantly higher, and subjective workload ratings were significantly lower for the graphic display and moderate information rate conditions. No differences were found for shared versus isolated operator/display settings except in the combined "worst case" condition of alphanumeric format and fast information rate, where a slight advantage was found for the shared operator/display setting. An exploratory analysis of group verbal and nonverbal communication behavior was also conducted. Isolated operator/display settings tended to produce more verbal activity and less gesturing than shared operator/display settings. Verbal activity was also found to be sensitive to information presentation rates.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Denise L. Wilson	22b. TELEPHONE (Include Area Code) (513) 255-7572 22c. OFFICE SYMBOL AAMRL/HED

DD Form 1473, JUN 86 Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

SUMMARY

The purpose of this report is to present research findings relating to the possible advantages of newly available technologies in the presentation of information to be shared by a command and control team. New technologies create potential for new work environments (group viewing of large screen displays), as well as new possibilities for information presentation (dynamic color graphics). The study described in this report utilized a complex dynamic task, requiring extensive interaction between team members, to investigate the utility of large screen group displays and dynamic color graphics in group decision-making. Both performance and subjective workload measures are reported.

PREFACE

This report was accomplished by the Crew Station Integration Branch, Human Engineering Division, of the Harry G. Armstrong Aerospace Medical Research Laboratory (AAMRL). It was performed under Work Unit 7184 27 03, "Multi-Operator Performance with Electronic Media." The research described in the report was supported by Systems Research Laboratories (SRL), Dayton, Ohio, under Contract Number F33615-85-C-0541. Mr. Robert Linhart is the contract monitor.

The authors gratefully acknowledge the support of the following people:

- Mr. Kevin Holloran--Subject Support
- Mr. Bill McGovern--Experimentation Support
- Mr. Charles Goodyear--Data Analysis
- Mr. Curtis Mayrand--Software Support

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Section 1
INTRODUCTION

BACKGROUND

The primary functions of a command, control, and communications (C^3) system are situation assessment and allocation of resources. Situation assessment involves the integration and evaluation of data from various sensor systems and other intelligence sources. Allocation of resources includes the alerting and queueing of specific weapons platforms, prioritization of targets, and weapons assignments. The accuracy and timeliness of these functions are critical to the flexibility and survivability of forces.

These situation assessment and resource allocation functions are performed by a command and control team. While each team member may have his own area of specialization, it is necessary for the team members to share information in order to fully assess the situation and make resource allocation decisions.

The purpose of this research is to explore the possible advantages of newly available technologies in the presentation of information to be shared by the command and control team. A technology of particular interest for this application is the large screen display for group viewing of information. The use of large screen displays acts to create a shared work environment, which is becoming common in modern command and control centers. Little data are available comparing the performance of teams in a shared work setting versus an isolated work setting. The literature reveals little previous work comparing performance of teams in such work settings. A few studies examined the utility of large screen displays using static tasks and photographic projection systems. The study described in this report utilized a complex dynamic task which required extensive interaction between members of three-person teams to achieve best performance.

Also of interest is the format for information presentation. Both dynamic color graphics versus monochromatic alphanumeric display formats were examined for their utility in the team performance task.

OVERVIEW

This report is organized as follows: Section 1 presents a review of the literature pertaining to team decision making research. Section 2 describes the Team Research Allocation Problem (TRAP) as it is implemented in this study, including the variations in graphic and alphanumeric display formats. Section 3 presents a review of the literature pertaining to the utility of graphic versus alphanumeric display formats to operator performance. Section 5 presents a discussion of workload measurement, including subjective techniques and Subjective Workload Assessment Technique (SWAT) in particular. Sections 6, 7, and 8 discuss the present study methodology, results, and conclusions. Appendix D, "Groups in TRAP: An Exploratory Analysis of Behavioral Dynamics," presents a pilot study assessing group verbal and nonverbal behaviors during team problem solving activities. This observational research was performed by A. Rodney Wellens and Tracy Vogler.

Section 2
TEAM DECISION MAKING RESEARCH

Situation assessment and resource allocation by a command and control team are complex processes which may be studied from a multitude of perspectives. Research in group behavior and decision making offers many theoretical orientations and empirical findings relevant to team decision making in C^3 settings. The following paragraphs first develop distinctions among command and control teams and other interacting groups. Group research is reviewed and the approach of the present research is examined.

In a recent review of team research, Dyer (1984) suggested that, "a team consists of (a) at least two people, who (b) are working toward a common goal/objective/mission, where (c) each person has been assigned specific roles or functions to perform, and where (d) completion of the mission requires some form of dependency among the group members..." (page 286). Dyer distinguishes teams from small groups where individuals' roles, functions, and interdependencies are usually less clearly defined.

Rather than focusing on teams, McGrath's (1984) typology of groups emphasizes the scope of activities, time duration, and composition of various groups. According to McGrath's typology, a command and control team would be considered a "standing crew" based on its limited scope of activities (they focus on the mission), relatively long duration (they may exist over months), and natural (real-world) versus concocted (existing to be studied) composition.

While it is desirable to study command and control teams in their operational environment, both practical and methodological difficulties of this approach have limited its application (Goldin and Thorndyke, 1980). Research using concocted groups has experimentally explored factors believed to affect team performance leading to empirical and theoretical development. Although the generalizability of the experimental approach is often questioned, clearly the preponderance of current empirical knowledge about team performance is based on ad hoc laboratory groups working on artificial tasks.

Although there are many excellent reviews of group performance (cf., Hackman and Morris, 1975; Hill, 1982), McGrath's (1984) recent book is comprehensive and integrative. His typologies of groups and tasks clearly illustrate the complexity of the field and the utility of systematically ordering empirical findings along meaningful conceptual dimensions. McGrath's circumplex model of group tasks, which distinguishes among eight task types based upon the dimensions of cooperation-competition and conceptual-behavioral, is a useful framework for understanding group performance. McGrath's (1984) viewpoint, that (a) there are many types of existing and concocted groups which may be studied; and (b) group performance depends largely upon task demands and the utilization of resources available to the group, including members' intellectual and interpersonal characteristics, is shared widely (Hackman and Morris, 1975; Hill, 1982; Shaw, 1976; Steiner, 1972; Shiflett, 1979). McGrath's (1984) research-based generalizations of group performance in various tasks, such as "truth supported wins" for intellectual tasks, "very strong majority wins" for decision-making tasks, and "minimal majority wins" for coalition tasks, lead him to conclude that performance groups are "pragmatically effective."

The importance of theoretical models, whether prescriptive or descriptive, is emphasized throughout the team and group process literature. Although no comprehensive team or small group theory exists (Dyer, 1984; McGrath, 1984; Shaw, 1976), there are a number of approaches to the study of small groups, as well as many midrange theories of group behavior, which suggest factors/processes likely to importantly affect command and control teams.

Contemporary social psychological theories of group behavior, for example, suggest that individuals in a group setting may (a) experience socially engendered arousal which may lead to performance enhancements or decrements (Geen and Bushman, 1978); (b) diffuse the impact of outside forces among other group members (Jackson, 1987); (c) engage in social comparison processes contributing to "group think" and group polarization effects (Geothals and Darley, 1987); (d) experience cognitive dissonance following collective action resulting in attitude change or diffusion of responsibility (Sande and Zanna, 1987); (e) demonstrate concern about self-presentation to other group members (Baumeister and Hutton, 1987); and (f)

become self-attentive, often leading to adherence to salient behavior standards (Mullen, 1987). The importance of social cognition (e.g., knowledge representation, person organization, confirmatory hypothesis testing) on behavior in groups has also been recognized (Pryor and Ostrom, 1987). Finally, Wegner (1987) has suggested that groups have a "transactive memory system," a contemporary group mind approach which emphasizes a group's unique information storage and retrieval mechanisms based on members' communications, roles, experiences, and expertise.

Shiflett's (1979) representational format for a general model of small group productivity provides a useful structure for theory development. Group productivity (P) is seen as the result of matrix operations representing the utilization or "transformation" (T) of resources (R). Thus, P represents outputs of a small group based upon multiplying matrix R, group resources (e.g., raw material, tools, knowledge, ability), by matrix T, various theoretically and empirically derived transformation rules (which may reflect role systems, task constraints, communication channels, and even personal characteristics). Shiflett's (1979) conceptualization, like some other approaches (e.g., Social Impact Theory; for a current summary, see Jackson, 1987), does not specify the actual processes or factors which underlie group performance, but instead suggests a meta-theoretical orientation for organizing and interpreting empirical results. For instance, Shiflett (1979) describes Steiner's (1972) disjunctive and conjunctive tasks as special cases of his general model. Different task types suggest alternative utilization (T) of individuals' abilities (R) resulting in unique functional relationships between group size and output (P). In an additive task (e.g., husking corn), each member's contributions enter into the group product in a simple additive manner. In a disjunctive task (e.g., solving an anagram), the most competent member's ability, and in a conjunctive task (mountain climbing), the least competent member's ability, determines group productivity. More recently, Stasser and Titus (1985) suggested that the process of biased information sampling in group decisions, where greater weight is given to information which is widely shared and supports initial preferences, could be characterized in Shiflett's terms as systematic tendencies to over-utilize (transform) certain resources (individuals' information).

While there are many theoretical views of groups, McGrath's overall characterization of group performance in decision making and problem solving as "pragmatically effective" parallels conclusions in cognitive psychology about individual performance. Although decision makers and problem solvers are often logical, efficient, and effective, both individuals and groups can demonstrate systematic errors and biases which leave them short of ideal models of performance (cf., Fiske and Taylor, 1984; Kahneman, Slovic, and Tverski, 1981; Steiner, 1972; Stasser and Titus, 1985).

Combining McGrath's group and task typologies with Dyer's team definition helps identify the problem area of interest in the present research. Concocted teams are studied under experimental conditions using a dynamic, intellective task demanding high intrateam coordination to achieve effective performance. Not unlike earlier research in group dynamics (Lanzetta and Roby, 1956, 1957; Leavitte, 1951; Morrissette, 1966; a generic, cognitively-based task is utilized. Developed with an appreciation of (a) the precision and control offered in experimental game research (Pruitt and Kimmel, 1977); and (b) the need to make social psychological research compelling and involving (Carlsmith, Ellsworth, and Aronson, 1976); a dynamic task environment is used which takes advantage of computer-interface technologies.

Section 3
TEAM RESOURCE ALLOCATION PROBLEM (TRAP)

The team resource allocation problem (TRAP) developed for this study is an extension of an experimental paradigm used by Pattipati, Kleinman, and Eprath (1982) to develop a Dynamic Decision Model (DDM) of human task selection in a multitask environment. In their research, individual subjects seated before a CRT observed rectangular boxes moving across the screen in several different rows. Each box represented a task which the subject could process by pressing a pushbutton corresponding to the row in which that task appeared on the CRT. Since many boxes crossed the screen concurrently in different rows, not all of the boxes (tasks) could be processed. Instead, the subject chose particular tasks for processing based on various manipulated task characteristics (i.e., reward value, required processing time). Experimental studies revealed that subjects' performance was sensitive to manipulations of task characteristics and consistent, overall, with the DDM.

Brown and Leupp (1985) extended the experimental paradigm from the study of individual task selection performance to the study of team performance. They found in their experimentation that three-person teams did respond in a meaningful manner to the properties of the TRAP.

The properties of the TRAP used in the present research are identical to that used by Brown and Leupp. The display format, however, has been modified to allow for evaluation of color graphic versus monochromatic alphanumeric formats. A thorough discussion of research literature pertaining to graphic versus alphanumeric display formatting is presented in the next section of this report.

The TRAP as currently implemented requires members of a three-person team (designated as persons A, B, and C) to work together to accumulate as many points as possible for the team by allocating their resources to the processing of selected targets. Each target requires the resources of a particular team member (A, B, or C) or some combination of the resources of team members (AB, AC, BC, or ABC) to process it, thus earning a specific

number of points for the team. Since more targets are available for processing than can possibly be processed by the team members' resources, team coordination is necessary in selecting optimum combinations of targets to obtain a high team score.

Targets differ not only with respect to the resources required for processing, but also with respect to their point value, which depends upon the color and shape designations of the target. While a target's value with relationship to its color designation is simple (red targets are worth more than blue targets), the target's value with relationship to its shape designation is more complex: one-resource targets (A, B, and C targets) are worth more if they are circles than if they are triangles; three-resource targets (those requiring simultaneous processing by all three team members' resources, ABC targets) are worth more if they are triangles than if they are circles; and two-resource targets (AB, AC, and BC targets) are unaffected by shape. Therefore, while target color is an easily generalized and interpretable attribute, target shape is interpretable only when analyzed along with the number of required resources. Table 1 presents all possible target parameter combinations and their respective point values.

TABLE 1. POSSIBLE TARGET PARAMETER COMBINATIONS AND RESPECTIVE POINT VALUES

One-Person Tasks	Point Value
Blue Triangle	1
Red Triangle	3
Blue Circle	3
Red Circle	5
Two-Person Tasks	
Blue Triangle	4
Blue Circle	4
Red Triangle	8
Red Circle	8
Three-Person Tasks	
Blue Circle	3
Red Circle	9
Blue Triangle	9
Red Triangle	15

In order to perform well, team members must evaluate the targets that are currently available to them and select for processing those target combinations which earn the team the greatest number of points. Success depends not only on choosing targets of high value, but also upon choosing them in a coordinated fashion. For example, person A may be able to individually process a red circle worth five points or work with person B to process a red triangle worth eight points. The optimal choice for the team would depend upon the additional target options available to the team members at that time. In fact, a simple model of decision making, which examines the available target options and chooses those which maximize the team score, serves as a useful standard for comparison with actual team performance.

The TRAP is represented in Figure 1 in its alphanumeric format. There are 11 rows in which targets may appear at random. The black squares in columns A, B, and C indicate which operators' resources are required for processing that target. The two-part target names (COLOR and SHAPE) are given in the next column followed by the time available to process that target. The last column indicates the status of the target if it is being processed or waiting for an operator's resources. To commit his resources to a particular target, the operator moves his marker, a white asterisk (*), to the corresponding black square. The marker is moved by pressing buttons marked "up" and "down" on the control box. The operator then pushes the start button. If the resources of only one team member are required for that target, processing automatically begins; the entire row is displayed in reverse video and the status column will display a numeric countdown of the processing time remaining for that target. However, if one or more additional team members' resources are required for processing the target, the status column will display a blinking "waiting" for this target. This means that at least one operator has committed his resources to this target and is waiting. Work will begin only when all the operators, whose resources are required for a target, have moved their markers to the target and pressed their start buttons. When this occurs, the entire row is displayed in reverse video, and a numeric countdown of the processing time remaining for that target is displayed in the status column.

ROW	A	B	C	TARGET	TIME	STATUS
1				RED CIRCLE	24	WORK
2				BLUE TRIANGLE	18	
3						
4				RED CIRCLE	7	
5				RED TRIANGLE	2	
6				BLUE TRIANGLE	15	
7				BLUE TRIANGLE	5	WORK
8				BLUE TRIANGLE	25	
9				RED TRIANGLE	20	
10				BLUE CIRCLE	13	
11				RED CIRCLE	29	
	A	B	C			
				FREE	ACCUMULATED POINTS	35
				WAIT		
	9	9	3	WORK		

Figure 1. Alphanumeric Representation of the TRAP Task

While an operator is processing a target or waiting at a target for another team member's resources, his marker becomes a circle. The marker may be moved to any row in preparation for the next target to which he may want to commit his resources. When his resources become free, the marker returns to its asterisk (*) form indicating that his resources are free to be committed to another target.

If an operator wishes to withdraw his resources from a target before processing is completed, he may do so by pressing the "reset" button on the control box. His marker will return to the asterisk form indicating that his resources are free. If other operators' resources were also involved in the processing of that target, they too will have to press their "reset" buttons to free their resources for another target. The team receives no points for targets which are not fully completed. If the team chooses, it may begin processing the same target again, but it will take the full ten time units (TUs) to complete the processing. The "reset" button may also be used by an operator to free resources which are "waiting" at a target.

Processing of each target requires ten TUs. A TU is some arbitrary number of seconds; the number of seconds is varied as a controlled parameter of the study. Each target is available for 30 TUs.

The countdown for each target starts at 30 TUs when the target is first displayed, and decreases at a constant rate as the 30 TUs available for processing the target elapse. When the countdown is at ten TUs, there is just enough time left to process the target. The target is deleted from the table when zero TUs remain.

The table in the lower left hand portion of the display indicates whether each team member's resources are free, waiting, or working. A black square indicates that a particular team member's resources are free, while a blinking "W" indicates that they are in a waiting state. When a particular team member's resources are working, a numeric countdown, in TUs, will indicate how much processing time remains until the team member will be finished with the current target.

The TRAP is represented in its graphic format in Figure 2. Again, there are 11 rows on which the targets appear at random. The targets move across the screen as the available time for processing expires. The time scale is indicated across the top of the display. The countdown for each target starts at 30 TUs when the target is at the left-most part of the screen, and decreases at a constant rate as the target moves to the right. When the target is at the end of the opportunity window (the black vertical line), the countdown is at 10 TUs. The target leaves the screen at 0 TUs.

The operator's marker is a green asterisk. While a particular operator's resources are working on a target or waiting for another team member's resources at a target, the marker turns red. When a team member's resources again become free, the marker returns to its green color.

If only one team member's resources are required for a target, work automatically begins when the start button is pressed, and the black square turns yellow. However, if the resources of one or more additional team members are required for processing the target, the black square turns

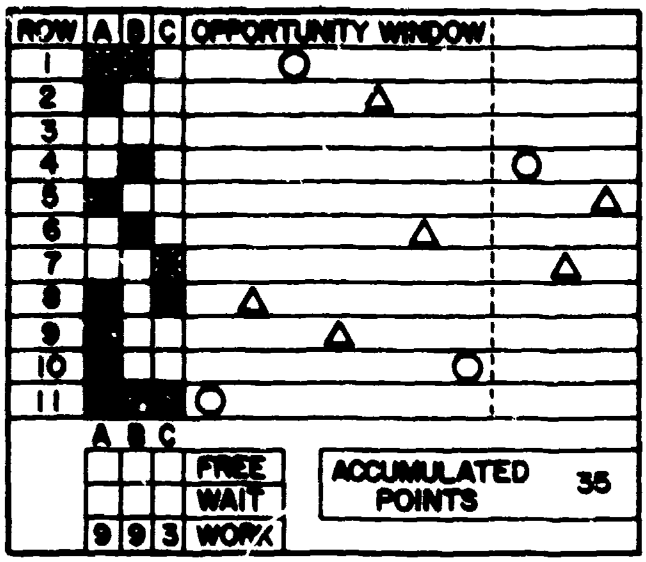


Figure 2. Graphic Representation of the TRAP Task

pink, indicating that certain resources are committed to that target but are waiting for other required resources. When the other required resources are committed to that target, the squares turn yellow indicating that work has begun.

When all the required resources have been committed to a target and work has begun, a black rectangle appears at that target's present position and extends through the 10 TUs required to process that target, thus indicating a time window for the processing of that target. The target moves through this time window as the processing proceeds. When the target moves out of the time window, processing of that target has been completed. In order for processing of a target to be completed before it leaves the screen, the processing must be started while the target is in the "opportunity window" (i.e. before it reaches the black vertical line at the 10 TU marker).

As in the alphanumeric format, if an operator wishes to withdraw his resources from processing a particular target before completing it, he may

do so by pressing the reset button on his control box. His marker will turn green, indicating that his resources are free. If the resources of others were committed to the processing of that target, they too will have to reset and select another target.

The table in the lower left hand portion of the display again indicates the status of each team member's resources. A black square indicates that a particular team member's resources are free. A blinking pink square indicates resources that are waiting. For working resources, the numeric countdown (in TUs) indicates how much processing time remains until these resources are free.

In summary, the object of the exercise is to accumulate as many points as possible as a team. Team members are encouraged to discuss alternatives in order to make optimum selection of the targets. As more targets are presented than the team can possibly process, combinations of targets should be selected which optimize team performance and total point count.

Within the context of the TRAP, various parameters can be varied experimentally to study their effects upon team decision-making. The parameters of interest in this particular study were time, work setting, and display format. How these parameters were varied will be explained in detail in Section 6 (Method).

Section 4

DISPLAY FORMATS: ALPHANUMERIC VERSUS GRAPHIC

Another facet of the shared versus own display question for facilitation of group decision making is the question of display formatting. Should the display format be alphanumeric or graphic? Should color be used?

Researchers in a variety of disciplines have been interested in the utility of graphics as decision aids and have investigated the ability of users to interpret accurately information presented graphically and make effective decisions. DeSanctis (1984) provides an extensive review of literature pertaining to how people process, interpret, and remember information displayed in graphical formats. More than one hundred sources were reviewed from areas of statistics, education, cartography, psychology, human factors, and management.

Statisticians were the first to investigate the utility of graphics. Most of their studies were concerned with the comparative effectiveness of various types of graphs and of graphics versus tables.

Human factors engineers have been concerned about the readability and comprehension accuracy of visual displays. Studies have examined the effects of brightness, contrast, texture, color, and shape coding upon the ability of the operator to quickly and accurately extract the needed information.

In the area of cartography, concern is with the design of maps for effective use by soldiers, pilots, sailors, etc. The communication effectiveness of map characteristics (dot symbols, color, and display size) has also been studied.

Cognitive psychologists have been interested in understanding the memory processes associated with various types of visual displays. They have studied the effects of visual displays for interpretation accuracy and comprehension of information presented.

Educators have been interested in the effects of visuals on human learning. Their studies have been concerned with the ability of visuals to complement narrative information.

In the area of marketing, the concern is with capturing interest. The use of illustrations or other methods of visual information presentation in advertising has been widely investigated.

In the management area, nearly all research in human use of graphics has focused on the tables versus graphs controversy. The concern is with presenting information in a format that is easily readable yet precise.

Across the various disciplines, interest in graphics has been associated with a limited set of possible outcomes:

- Graphics might permit more rapid presentation and assimilation of data.
- Graphics might serve to enhance the meaningfulness of presented data.
- Graphics might serve as a data reduction tool.
- Graphic presentation of information might result in improved decision quality.
- Graphic presentation of information might afford greater decision speed.

In their investigation of the utility of graphic versus alphanumeric information presentation, researchers have most often studied interpretation accuracy, problem comprehension, and task performance. Some researchers have examined user attitudes toward graphics, particularly preference for graphics versus other support tools. Psychologists and visual information processing researchers have studied reaction time (speed of response) using graphic versus other display formats. Their rationale being that design

criterion for visual stimuli should be that they require minimum cognitive effort for the reader.

The studies reviewed employed various measures of interpretation accuracy, information comprehension, and data recall. Research results were found to vary as a function of the task as well as the type of graphic employed. DeSanctis (1984) concludes that consistent patterns in the results of existing research are difficult to identify. She further postulates that the information obtained from any display most likely interacts with variables such as task complexity, task content, degree of task structure, or the quality of available decision alternatives, all of which are known to affect decision making.

A few studies not included in the DeSanctis literature review do lend support to the utility of graphic display formats for high workload, high time-stress, complex decision tasks. Dickson, DeSanctis, and McBride (1985) performed three experiments which compared decision-supporting effectiveness of two media (tables versus graphs) in terms of interpretation accuracy and decision quality. They employed increasing task structure and task complexity across the three experiments.

For the first experiment, task content was an accounting setting very familiar to all subjects; the task is described as being relatively simple. Task complexity was considered low; subjects had only one variable to process at a time. The task was highly structured; the subjects were provided with a step by step procedure.

Interpretation accuracy and decision quality scores were not significantly different for the two groups. On a 7-point Likert scale, the graphic group found the reports to be more difficult than the tabular group. The experimenters' interpretation of the results of this experiment is that for simple financial reporting, use of graphical displays will neither hinder nor improve interpretation accuracy or decision quality over that observed with traditional tabular reports. They observe that users appear to prefer a report format with which they are familiar.

Task content for the second experiment was somewhat more sophisticated and involved forecasting demand for a product. The task was less structured; there were no guidelines. The task was more complex; subjects were required to process two variables at a time.

Interpretation accuracy was found to be better with the graphic format, although the effect was not statistically significant. The graphic group produced statistically better decision quality. The graphic group also perceived the task to be easier than tabular group. The experimenters' interpretation of these results is that data presentation using graphs enhanced decision quality. These results imply that it may not be appropriate to generalize about the superiority of a particular method of data presentation across task environments.

The third experiment involved unfamiliar task content. The task was highly unstructured; there were no guidelines/no historical basis for performing the task. This task was of increased complexity in that it required simultaneous consideration of a number of variables. In addition to tabular versus graphic formats, complete presentation versus subset presentation, and recall versus lookup were studied.

In terms of "getting the message," there were no significant differences for display format or interactions involving format. For interpretation accuracy, the complete versus subset x format interaction was statistically significant. This suggests that if there is much information and many questions to answer, then graphics are better.

Stollings (1982) evaluated the relative effectiveness of three display format types in the context of CRT weapons displays used in advanced fighter cockpits. The three format types were:

- <u>Alphanumeric</u>: Complete words and numbers.

- <u>Black and White Graphic</u>: Pictorial example of aircraft and bomb load status.

• <u>Color Graphic</u>: Added color to the graphic representation.

The subjects were 24 highly experienced fighter pilots.

Stollings found that when exposure durations were controlled by the subjects, the time required to answer questions about display status was not affected by display format. Complex questions, however, required more time than simple questions.

Display format differences were revealed in error frequencies under controlled exposure durations. Color graphics formats were found to be superior to alphanumeric and black and white formats when (a) short exposure durations were used, (b) complex questions were asked, and (c) questions were asked after the subjects had viewed the display. Stollings concluded that display format makes a difference in high workload, high time-stress, situations which require complex processing. These findings seem to be in agreement with those of Dickson et al. who found better interpretation accuracy and decision quality for graphics format in a complex task but no difference for display format in a simple task.

Scott and Wickens (1983) investigated the utility of graphics in a decision task associated with a tactical battlefield scenario. They had subjects perform a probabilistic information integration task in which multiple cues were presented in either digital or analog-graphical form. The cues were in favor of either of two competing hypotheses and varied in their reliability and their diagnostic value in choosing between the hypotheses.

The subjects were asked to imagine themselves as commanders of a defensive unit preparing for attack from either north or south. The subjects received a series of intelligence cues, each identified by source, diagnosticity (0 to 1.0), and reliability (0 to 1.0). Based on these cues, the subjects were required to determine from which direction attack would take place and give a confidence rating. For half of the problems, the cues were presented in a verbal format. For the other half, a graphic-spatial format was employed. Reliability and diagnosticity were defined by the width and height of a rectangle respectively. The total worth or balance

of a cue then = reliability x diagnosticity = area of the rectangle. Other independent variables were problem size (6, 7, or 10 cues), net evidence (weak, medium, or strong), and presentation speed (5 cue/second or 3 cue/second).

The spatial display was found to yield reliably more accurate judgments than did the verbal display. In using the graphic format, no mental multiplication was required; the subject had only to integrate the areas of the rectangles associated with the various cues favoring one or the other of the two hypotheses. The complexity of the task was, therefore, reduced when the graphic format was used. These findings are in agreement with those of Dickson et al. and Stollings who found graphic formats to be advantageous in complex problem-solving situations.

Schwartz and Howell (1985) sought to determine whether and, if so, how performance varies as a function of the manner in which a progressively unfolding decision problem is displayed over time. Their optimal-stopping decision problem involves a sequence of decisions. Each stage requires a choice between sampling additional information or selecting a terminal action. The cost and payoff functions are such that the operator can reduce the uncertainty of the decision problem only at the expense of the marginal utility of the ultimate choice. This type of problem represents a common real-world decision situation in which time is important. The experimenters sought to determine whether sampling tendencies, decision quality, and/or decision efficiency are affected by display format.

The subjects were required to monitor a series of simulated hurricanes in their advance toward a heavily populated target area. The storm paths were represented either graphically or alphanumerically with reference to an 8 (longitude) x 7 (latitude) grid. Storms always progressed from left to right in stepwise fashion. Uncertainty in the storm's path was introduced by building a certain probability of unit latitudinal movement into each advance. After each advance, subjects were required to make one of three responses: (1) stay (minimize potential losses associated with decision delay as in committing to intensified protective measures), (2) evacuate (total abandonment of the city), or (3) wait (postpone any terminal

action). The subject's task was to decide when to stop gathering information and, at that point, decide which action to take based upon the storm's current location and the subject's expectation of costs associated with the various options.

The first experiment was self-paced. No significant difference was found for display format. This finding is consistent with those of Stollings' self paced display presentation experiment.

A second experiment was performed in which three levels of time stress were introduced. Decision accuracy was found to be consistently superior using the graphic format, particularly in the more stressful pacing conditions. Again, the findings are consistent with those of Stollings for time-stressed tasks and with those of Scott and Wickens for the probabilistic information integration task. The authors hypothesize that graphic representation of an evolving decision problem encourages a more holistic processing strategy while numeric representation encourages a more analytic strategy.

The studies reviewed offer limited evidence in favor of graphic display formats over alphanumeric formats in high workload, high time-stress decision tasks which require complex processing. Based on these observations, one might expect the graphic display format to facilitate better TRAP task performance than would the alphanumeric display format under time-stressed conditions.

Section 5
WORKLOAD

OVERVIEW

A review of the literature pertaining to workload reveals that there is no single definition of workload and no universal measure of it. While the exact definition of mental workload is in dispute, generally the term refers to the amount of an operator's capacity that is expended in the performance of a given task within a certain system. The amount of capacity used is a function of the individual, systems demands, and task demands (Reid et al., 1984). Several investigators reviewed have conceptualized workload as a multidimensional construct (Johannsen et al., 1979; Sanders, 1979; Williges and Wierwille, 1979). Some combination of factors relating to the tasks to be performed, factors associated with environment, and internal subject factors are the components of most definitions (Reid, 1982). Johannsen et al. (1979) have suggested that workload is composed of behavioral, performance, physiological, and subjective components. Sanders (1979) had defined workload as a composite concept which reflects the end result of various contributing factors related to task, internal operator dispositions, and the state of operator practice.

Moray (1982) and Johannsen et al. (1979) discussed the distinction between physical and mental workload. As human-machine systems become more complex and automatic control more sophisticated, the operator may appear to do less. This reduction in physical activity, however, is not necessarily indicative of a corresponding reduction in workload experienced by the operator. Mental information processing and other dimensions of mental workload are exhausting and stressful and do utilize the body's physical energies. The proportion of physical to mental energy load that is required by a given task is difficult to quantify; "all tasks have some of each component in their total contribution to the human operator workload" (Johannsen et al., 1979). The predominance of mental tasks is responsible for unprecedented increases in the amount of workload experienced by the

operator which may compromise the performance of the entire human/machine system (Reid et al., 1984).

Operationally, workload has been defined objectively and subjectively. Moray (1980) made a distinction between imposed mental load and subjective mental load. He defined imposed mental load as the load demanded by the task and measured by task parameters. Subjective mental load he defined as the load perceived or experienced by the operator.

Workload is believed to represent a multidimensional construct which includes physical, physiological, cognitive, emotional, subjective, and psychomotor components. It reflects the interaction of elements such as task and system demands (environmental, situational, and procedural) operator characteristics such as processing capabilities, effort, subjective performance criteria, information processing strategies, and training and experience.

WORKLOAD MEASUREMENT

Just as there is no single agreed upon definition of workload, there is no universally accepted metric of workload. Considerable scientific effort has been directed toward defining workload and developing methods for measuring it (Reid et al., 1984). However, most researchers do agree upon a set of characteristics that any measurement technique should possess. Any measurement technique should have face validity; it should seem an intuitively appropriate measure. It should be sensitive to the entire range of specific human performance from underload, where almost none of the operator's capacity is being employed, to overload, where all of the operator's capacity is being utilized and more is needed. The measure should be non-intrusive or at least reasonably unintrusive; measures which interfere with the operator's normal activities may yield invalid results. Generalizability is an important attribute; the measure should yield stable reliable results between and within people and situations.

There are three major categories of workload measurement techniques: physiological, behavioral or performance, and subjective. Physiological

methods involve the measurement of one or more variables related to the human physiological process. The underlying assumption is that as operator workload changes, involuntary changes take place in the physiological processes of the human body (body chemistry, nervous system activity, circulatory or respiratory activity, etc.) (Wierwille, 1979). O'Donnell (1979) gives a complete discussion of the various physiological measurement techniques.

The logic underlying behavior/performance based measures is that external behavior reflects internal events and processes. It seems logical to suppose that an operator is beginning to exceed his ability to process information and/or generate appropriate responses when he begins to make errors. The two major classes of performance-based measures are primary or single task measures and secondary or dual task measures. In the former, performance is measured for one or more tasks performed separately; in the latter, two tasks are performed simultaneously and performance on the lower priority task is taken as an index of the amount of mental capacity not required for the primary task. Both methods are based on the assumption that there is an upper limit to the amount of effort that can be exerted to meet task demands, and that decrements in performance will begin to appear as this upper limit is approached. Single task measures are discussed in detail by Shingledecker, Crabtree, and Acton (1982) and secondary task measures by Eggemeier (1981).

The use of subjective measures of workload is based on the rationale that if an operator feels loaded and effortful, he is loaded and effortful, regardless of what performance measures might demonstrate (Johannsen et al., 1979). Johannsen et al. have suggested that prior to performance breakdown, the operator might be working harder to avoid such decrements, and that subjective feelings could be used as an indicant of the additional effort which precedes degraded performance. Gartner and Murphy (1976) have indicated that when subjective impressions of workload are accepted, the operator's direct perception or estimation of his feelings, exertion, or condition may provide the most sensitive and reliable indicators of workload. Moray (1980) has pointed out that an objectively easy task may be experienced as difficult due to factors such as fatigue or motivation.

Given appropriate instructions and a balance between speed and accuracy, an objectively difficult task may be experienced as less effortful or difficult. In addition to their theoretical importance, subjective techniques have a number of characteristics which contribute to their potential utility as measures of operator workload (Eggemeier, 1981). They are relatively easy to implement and support when compared with many physiological and performance-based measures. Subjective measures minimize instrumentation requirements and, therefore, might be more easily implemented in an operational environment. If implemented correctly, the measures can be relatively nonintrusive and should not disturb primary task performance. If the general factors that contribute to workload can be identified, subjective measures could be applicable across a wide range of situations, while performance-based techniques are, by necessity, situation specific. A variety of subjective assessment techniques have been reported in the literature. Daryanian (1980) used a Thurstonian paired-comparison procedure to generate an interval scale of workload related to a multicomponent decision task. Hicks and Wierwille (1979) applied the method of equal appearing intervals to generate rating scale responses. This method successfully discriminated a number of workload conditions in a driving simulator. Borg (1978) has reviewed a program which made use of magnitude estimation techniques and category scales to develop indices of perceived difficulty in a group of physical and cognitive tasks. The program explored the relationship between perceived difficulty and task characteristics for several cognitive tasks. High correlations were obtained between subjective and objective measures of difficulty, supporting the capability of subjective ratings to reflect objective levels of task difficulty.

Wierwille and Connor (1983) compared the sensitivity and intrusion of 20 pilot workload assessment techniques using a psychomotor loading task in a three degree-of-freedom moving-base aircraft simulator. Of the 20 pilot workload assessment techniques studied, only two were opinion measures, and both proved sensitive. Wierwille and Connor (1983) conclude, "This suggests that well designed rating scales are among the best techniques for evaluating psychomotor load."

SUBJECTIVE WORKLOAD ASSESSMENT TECHNIQUE (SWAT)

Most subjective assessment techniques have been developed for a particular application and are not easily generalizable. Usually there are no data establishing that they validly and reliably measure workload. Readily available measures like the Cooper-Harper and the Systems Operability Measurement Algorithm (SOMA) scales assess handling qualities and systems operability, respectively. These measures, therefore, focus on systems evaluation with workload as a component (Reid, Eggemeier, and Shingledecker, 1982). "The appropriateness of using these or similar measures without considerable alteration and/or supporting data is questionable" (Reid, Eggemeier, Nygren, 1982). As an answer to the need for a systematically developed and validated subjective measure that is generally applicable and widely accepted, Subjective Workload Assessment Technique (SWAT) has been developed by the Armstrong Aerospace Medical Research Laboratory as a candidate generalized procedure for scaling pilot mental workload (Reid et al., 1981). SWAT uses a psychometric technique known as conjoint measurement to construct interval level workload scales from ordinal rankings of combinations of levels on three contributory dimensions.

Conjoint measurement (Coombs, Dawes, and Tversky, 1970; Krantz and Tversky, 1971) is a technique by which the joint effects of several factors are investigated and the rule or composition principle that relates the factors to one another is extracted from the data. A major advantage of this procedure is that only the ordinal aspects of the data are required for the production of an interval level scale which represents the joint effect of the factors.

SWAT distinguishes three levels for each of three dimensions: time load, mental effort load, and psychological stress load. These are adaptations of the categories defined by Sheridan and Simpson (1979). Time load refers to how much time is available for an operator to perform a task; this includes both overall time and task pacing. Can the person take as much time as he wishes, or must he hurry? Are there interruptions caused by overlapping task demands? (Reid, 1982). Mental effort load refers to the amount of attentional capacity or effort required without regard to the

amount of time available or task pacing. This might involve performing computations, reasoning, problem solving, or making decisions (Reid, 1982). Stress load refers to anything that makes the task more difficult by producing anxiety, frustration, and confusion; this includes such things as fatigue, stress, and fear, as well as physical stressors like vibration, G loading, and heat. The effects of the physical stressors that are included in this dimension are effects that occur prior to direct interference with task performance (Reid, Eggemeier, and Nygren, 1982). The primary assumption of SWAT is that workload can be adequately represented by the combination of these three dimensions.

SWAT is a two step process consisting of a scale development phase and an event scoring phase. These are two distinct events which occur at different times. During the scale development phase, the data necessary to develop a workload scale are obtained from a group of subjects. At the event scoring phase, the subjects rate the workload associated with a particular task and/or mission segment.

The three dimensions (time, effort, stress) taken in all possible combinations yield a 27-cell three-dimensional matrix to represent workload. To develop the scale, each subject rank orders the 27 combinations of verbal descriptors according to his own perception of the workload represented by each combination. The card sort data are then tested for agreement among subjects. If sufficient agreement exists to produce a Kendall's coefficient of concordance of 0.75 or greater, then group data are used to develop the SWAT scale. If agreement is low, then homogeneous subgroupings are accomplished by correlating individual subject rankings with rankings of what have been termed SWAT model prototypes. The prototypes are based upon the assumptions that (1) time, effort, and stress are combined according to an additive rule, and (2) the prototype rankings represent perfect data with consistent weightings assigned to each of the three dimensions. All possible combinations of weightings yield six prototype rankings. The individual subject rankings are then correlated with the six prototype rankings. The highest correlations identify which dimension or dimensions the subject considers to have the highest weight. "Considerable

gain in subjects' level of agreement is realized by placing them into groups" (Reid, Eggemeier, and Nygren, 1982).

The results of the ranking procedure are then used to develop an overall interval workload scale which represents the joint effect of the three dimensions. The composition rule for the ordered data is defined through a series of axiom tests. When the appropriate rule has been identified, the scaling transformation is computed. The transformation fits the data to the defined model while maintaining the order inherent in the original data (Reid et al., 1982). If the prototype procedure has identified different groups, the scores of the various group scales are standardized into a final SWAT scale.

The event scoring phase is an implementation of the scale as a dependent variable. This is accomplished, as with other rating procedures, by analyzing the tasks or mission scenario to determine what ratings are needed, what ratings are possible given the scenario, and when the ratings should be obtained. A major positive attribute of SWAT is the simplicity of the event scoring procedure. The events are rated using the same descriptors previously used for scale development. Asked to provide a SWAT rating for a particular event, a pilot would assign either a 1, 2, or 3 to each of the three dimensions of time load, effort load, and stress load experienced during that event. The numbers for each level of the three dimensions are defined as in the scale development phase, and these definitions are supplied to the pilot for reference. These three ratings correspond to one of the combinations created in the ordering procedure for scale development. The scale value computed for this particular combination of the three factors is the subjective workload score assigned to the event. Therefore, if the investigation were to compare workload associated with various system architecture options, SWAT ratings would be obtained for selected mission segments using each option under investigation. The workload scale values (for example, 85 for option A and 76 for option B for a particular mission segment) would then become the dependent variable for evaluating the workload associated with the two options.

An important aspect of the SWAT development program has been determining the sensitivity of the technique to workload variations in a number of different types of tasks. Investigations reported by Reid, Shingledecker, and Eggemeier (1981) have demonstrated that subjective ratings derived from the SWAT procedure were sensitive to variations in the difficulty of a simulated aircrew radio communications task and a critical tracking task. Reid, Eggemeier, and Shingledecker (1982) found that SWAT ratings were sensitive to a number of difficulty manipulations on a high fidelity simulation of an air-to-air combat task.

To further explore the general applicability of the SWAT technique, Eggemeier et al. (1982) examined its sensitivity to variations in the rate of stimulus presentation and levels of information load in a sequential short term memory task. The memory update task was intended to be representative of portions of an air traffic controller's task. The SWAT procedure was found to be extremely sensitive to variations in presentation rate and was more sensitive than memory error in all but the most difficult condition involving variations in levels of information load. SWAT ratings were found to have greatest relative sensitivity at the lowest levels of workload. The data are consistent with the expectation that at lower levels of workload (nonoverload situations), primary task measures will be relatively insensitive to variations in load compared to more sensitive workload assessment techniques. Primary task measures can be expected to increase in sensitivity as the threshold for performance breakdown is approached. This study demonstrated the capability of SWAT to provide a sensitive index of variations in workload at levels below the threshold for performance breakdown.

Further analysis of the above study (Eggemeier, McGhee, and Reid, 1983) indicated that the three component scales of the SWAT procedure demonstrated complementary sensitivity patterns to variations in the rate of stimulus presentation and levels of information load. Eggemeier et al. interpret these sensitivity patterns as supporting the inclusion of time, effort, and stress scales in SWAT. Eggemeier et al. also suggest that subscale analyses may be used to provide more diagnostic information regarding the potential sources of high loads which had been identified.

Eggemeier and Stadler (1984) found SWAT to be sensitive to two of the three difficulty manipulations in a spatial short term memory task. Subjects were required to retain a histogram pattern over a short retention interval and then indicate whether a subsequently presented histogram matched the original or "target" pattern. Task difficulty was manipulated by varying the complexity and spatial orientation of the "target" histogram and by varying the length of the retention interval. The SWAT procedure was found to be sensitive to both target complexity and retention interval. SWAT demonstrated greater sensitivity than either errors or reaction time which were the primary task measures.

While there is no single definition of workload, it is generally agreed that workload is a multidimensional construct made up of physical, physiological, cognitive, emotional, subjective, and psychomotor components. This multidimensionality of workload supports the inclusion of subjective measures of workload as well as measures of performance of the TRAP task.

Section 6
METHOD

EXPERIMENTAL VARIABLES

The experimental conditions studied in the present research were type of work setting, display format, and information rate. Multioperator/display setting was the variable of primary interest. Two conditions were studied: (1) team members seated together in the same room and sharing a single large screen display, and (2) team members seated in distinctly isolated cubicles, each with an individual CRT display. In both cases, the team members worked with individual control boxes and communicated with one another by means of microphone and headphones. Two display formats were implemented: (1) monochromatic alphanumeric (Figure 1) and (2) color graphic (Figure 2). Both display formats were dynamically changing as targets were selected for processing, new targets became available, and time expired for others. Information rate was either (1) moderate, or (2) fast. This refers to the speed at which new targets became available and their required processing times. These rates were determined through pilot testing such that the moderate rate was subjectively comfortable but did not allow "dead time" while the fast rate produced a subjective feeling of time stress. For the moderate information rate condition, a new target appeared on the screen every 5.45 seconds and remained on the screen for 60 seconds of which 20 seconds were required for processing. For the fast information rate condition, a new target appeared on the screen every 2.73 seconds and remained for 30 seconds of which 10 seconds were required for processing. A TU, then, was equivalent to 2 seconds or 1 second for the moderate and fast information rates, respectively.

EQUIPMENT

The alphanumeric and graphic formats were controlled by a Silicon Graphics Model 2400 graphics display processor. Control response buttons on the custom built response boxes interfaced with the computer via an analog-to-digital converter.

The large screen display was generated by a General Electric Model PJ-5150 professional large screen projector. It is a full color, oil film light valve projection system, and has a minimum usable resolution of 750 horizontal x 650 vertical, with a 1023 line, 60 frames per second scan standard. The image was rear projected onto a Phoenix Inc Model XX screen. Large display image size was 89 cm high by 126 cm wide. Character height was 1.9 cm, subtending an angle of 21.5 minutes of arc at the 305 cm viewing distance.

The three individual displays were Conrac Model 7211, 33 cm diagonal, full color raster scan CRTs with resolution of 921 horizontal x 739 vertical pixels, and 40 MHz video bandwidth. Character height on the CRTs was .48 cm, subtending an angle of 22.9 minutes of arc at the 46 cm viewing distance.

SUBJECTS

The subjects were 48 paid volunteers between the ages of 18 and 30 years. All subjects exhibited 6/6 normal or corrected visual acuity and normal color vision. Sixteen teams (three persons per team) participated in the present study in a full factorial within teams design.

PROCEDURE

Following instructions and a short test of point value understanding, each team completed two reduced pace demonstration trials during which they were encouraged to ask questions and practice using their response boxes. Four additional practice trials concluded the training session.

Each of the four test sessions consisted of eight trials. Each trial consisted of 44 targets: 11 blue triangles, 11 red triangles, 11 blue circles, and 11 red circles. Of the 11 targets of a given color and shape, six were one-person targets (two A, two B, and two C targets), three were two-person targets (one each of AB, AC and BC targets), and two were three-person targets (two ABC targets). In this way, one third of the targets which a particular participant could be involved in processing were

one-person targets, one third were two-person targets, and one third were three-person targets. These targets were surrounded by a buffer of ten targets at the beginning of the trial and six targets at the end which were not included in the data analysis.

Sixteen unique random sequences of the 44 targets were created. Four of the random sequences were utilized within each session; the same random sequence was presented once as a fast information rate trial and once as a moderate information rate trial. Within each session, the trials alternated between moderate and fast information rate, so that four trials were run at the moderate information rate and four at the fast information rate. The operator setting and display format variables were varied between sessions with the order of presentation controlled by two Latin squares. (The Latin squares also controlled the order of presentation of the 16 unique random sequences of the 44 targets.)

The primary dependent variable was team score. This is the number of points per trial accumulated by the teams minus those points earned for tasks in either the beginning or ending buffer zone. A more detailed analysis of target selection as a function of shape, color, and number of persons required confirmed that teams were indeed sensitive to the combinations of parameters which determine target point value. These analyses provide a more detailed understanding of team members ability to optimize target selection across the various experimental conditions and will be discussed in the following section.

In addition to the performance measure associated with team score, subjective workload ratings were also collected by means of the SWAT discussed in the preceding section of this report.

The rationale behind the collection of both performance and subjective workload measures was that teams or individuals may feel more loaded or effortful under some conditions yet maintain a constant level of performance by exerting more effort. Performance measures, then, are not sensitive to these differences in workload. SWAT has been shown to be sensitive

to variations in workload that occur below the threshold for performance breakdown.

SWAT ratings were collected after each trial. The descriptors for each of the three levels of time stress, mental effort, and psychological stress appeared on the screen and the subject pressed a button designated 1, 2, or 3 on the control box to indicate the selected level for each of the three dimensions.

Section 7
RESULTS

SWAT SCALE DEVELOPMENT

As was mentioned in Section 5, the SWAT procedure consists of two stages: scale development and event scoring. The analysis of the card sort data and development of the interval level unidimensional workload scale for each subject is discussed first. Following sections then discuss the analysis of the event scoring.

Through the process of conjoint measurement, the card sort data are converted to an interval level unidimensional workload scale for each subject. The 27 cards, which the subjects were asked to sort according to their individual perception of workload, represented all possible combinations (one combination per card) of the three levels of the three SWAT dimensions ($3^3 = 27$). In almost all cases a 1,1,1 triplet is perceived to represent the lowest level of workload and a 3,3,3 triplet, the highest level of workload. It is the arrangement of the intervening combinations that reflects the subjects' individual perceptions of workload.

The arrangement of the combinations usually reveals that a subject is dominated by either time, effort, or stress in his perception of workload. For example, a particular subject who is acutely sensitive to time may arrange the 27 combinations to reflect time as an "outer loop." His arrangement would tend to reflect a perception of higher levels of workload with moderate and low levels of time stress than with corresponding (or perhaps higher) levels of the other two dimensions. A group of subjects who are dominated by the same dimension are said to be of the same prototype. Although the conjoint scaling procedure produces a separate workload scale for each individual subject, the workload scales produced for subjects of the same prototype are usually highly correlated. In this case, it is preferable to compute a group workload scale for these subjects. This is because individual subjects generally do not give perfect data, and the scaling solution for the individual does not usually fit the model as well as a scaling solution for a group. When a group scale is calculated,

individual subject errors tend to average out, resulting in a group scale which better represents the underlying construct (Reid et al., 1982).

Of the 48 subjects who completed the card sort and participated in this study (16 teams of three subjects each), 11 were of the time prototype (Kendall's Coefficient of Concordance = 0.84), nine were of the effort prototype (Kendall's Coefficient of Concordance = 0.78), and 28 were of the stress prototype (Kendall's Coefficient of Concordance = 0.83). Therefore, three prototypical group scales were used in processing the SWAT event scores for analysis. These scales are shown in Table 2. The table is used to convert the reported triplets into the interval level unidimensional workload scales. The left-most column contains all possible triplets; the other three columns represent the three scale prototypes. For each raw score triplet, the workload value is read from the same row in the column of the appropriate prototype. For example, a 2,2,2 reported by a stress dominated individual would correspond to a workload value of 41.5, while the same 2,2,2 triplet reported by a time dominated individual would correspond to a workload value of 38.7.

TRAP PERFORMANCE AND SWAT RATINGS

Separate Analyses of Variance (ANOVAs) were performed for team score and SWAT ratings. The results of these ANOVAS are given in Table 3. Team scores were found to be significantly higher ($p < 0.01$) and SWAT ratings significantly lower ($p < 0.01$) when using the graphic display format than when using the alphanumeric format. The mean team scores pooled over operator/display setting and information rate were 101.50 and 89.74 for the graphic and alphanumeric formats, respectively; while the mean SWAT ratings were 35.38 and 41.53 for graphic and alphanumeric formats, respectively. The effect of information rate upon team scores and SWAT ratings was also found to be significant ($p < 0.01$). Team scores were significantly higher and SWAT ratings significantly lower for the moderate information rate than for the fast information rate. The mean team scores, pooled over display format and operator/display setting, were 103.38 for the moderate information rate and 87.86 for the fast information rate; while SWAT ratings were 33.00 and 43.90 for moderate and fast information rates, respectively.

TABLE 2. PROTOTYPICAL GROUP WORKLOAD SCALES

Prototype	Time	Effort	Stress
1,1,1	0.0	0.0	0.0
1,1,2	10.2	13.1	21.5
1,1,3	24.4	26.9	52.1
1,2,1	7.3	25.9	11.4
1,2,2	17.6	39.0	32.9
1,2,3	31.7	52.9	63.5
1,3,1	25.9	43.8	24.1
1,3,2	36.1	56.9	45.6
1,3,3	50.2	70.8	76.2
2,1,1	21.1	10.6	8.6
2,1,2	31.3	23.7	30.1
2,1,3	45.5	37.6	60.7
2,2,1	28.4	36.6	20.0
2,2,2	38.7	49.6	41.5
2,2,3	52.8	63.5	72.1
2,3,1	47.0	54.5	32.7
2,3,2	57.2	67.5	54.2
2,3,3	71.3	81.4	84.8
3,1,1	49.8	29.2	23.8
3,1,2	60.0	42.3	45.3
3,1,3	74.1	56.2	75.9
3,2,1	57.1	55.2	35.2
3,2,2	67.3	68.2	56.7
3,2,3	81.5	82.1	87.3
3,3,1	75.6	73.1	47.9
3,3,2	85.9	86.1	69.4
3,3,3	100.0	100.0	100.0

TABLE 3. SUMMARY OF ANOVA RESULTS

Source	Performance	SWAT
Format (F)	$p < 0.01$	$p < 0.01$
Setting (S)	NS	NS
Info Rate (I)	$p < 0.01$	$p < 0.01$
F x S	NS	NS
F x I	$p < 0.05$	NS
S x I	NS	NS
F x S x I	$p < 0.05$	NS

Analyses of Covariance (ANCOVAs) were performed to determine whether correlations existed between SWAT ratings and performance scores. The analyses used SWAT rating as the dependent variable with performance score as a covariate. Depending on the analysis, team, format, operator/display setting, and information rate were factors. Results of the ANCOVAS are presented in Table 4.

TABLE 4. SUMMARY OF ANCOVA RESULTS

Factors	Slope	Partial R	p Value
Team	-0.39	-0.57	0.0001
Team, Format	-0.37	-0.51	0.0001
Team, Setting	-0.40	-0.58	0.0001
Team, Info Rate	-0.25	-0.37	0.0001

The slopes did not vary significantly among the teams, formats, operator/display settings or information rates ($p > 0.05$). For example, adjusting for differences in teams, an increase of 10 in performance score corresponded with a decrease of 3.9 in SWAT rating regardless of how format, operator/display setting, or information rate were manipulated.

The display format x information rate interaction was also found to have a significant effect upon team score ($p < 0.05$). Examination of the data reveal a greater advantage of graphic format over alphanumeric format in

the fast information rate condition (14.32 points) than in the moderate information rate condition (9.19 points).

Although the main effect of operator/display setting had no significant effect upon team score ($p > 0.05$), a three way interaction involving operator/display setting with information rate and display format was significant ($p < 0.05$). This interaction is represented in Figure 3 and suggests an advantage for the group setting over the isolated setting only for the combined worst case conditions of alphanumeric format and fast information rate.

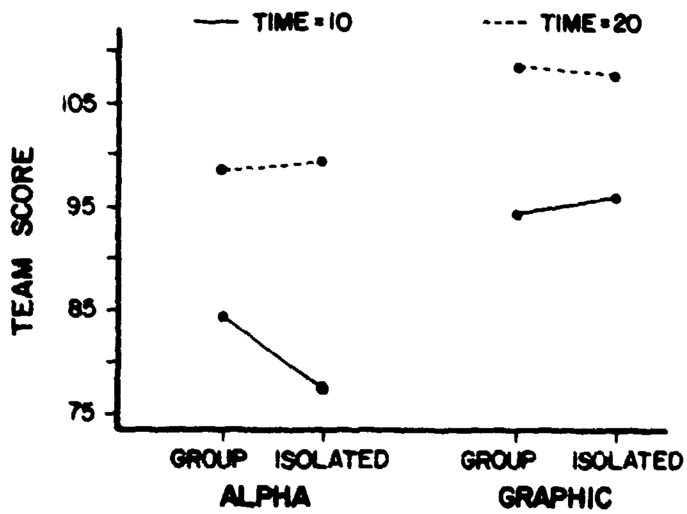


Figure 3. Team Score as a Function of Display Format, Information Rate, and Operator/Display Setting

DETAILED ANALYSIS OF TEAM PERFORMANCE

To allow a more detailed analysis of team performance, the TRAP used a highly structured set of targets. Specifically, the targets available for processing on each trial represented a complete 3 (resources: one, two, three) x 2 (shape: circle, triangle) x 2 (color: blue, red) factorial design. Therefore, the proportion of processed targets of each type averaged across four trials was analyzed as a function of the major independent variables of the study (format, setting, and information rate), as well as the target variables listed above. Together, this resulted in a completely

crossed 2 x 2 x 2 x 3 x 2 within teams analysis of variance of proportion of targets completed.

The analysis tested six main effects, 15 two-way interactions, 20 three-way interactions, 15 four-way interactions, six five-way interactions, and one six-way interaction. Of these 63 effects, 11 were significant at $p < 0.01$ and three more at $p < 0.05$. These significant results are presented in Table 5.

Five of the six main effects were statistically significant. Teams processed a higher proportion of targets when: (1) format was graphic ($M = 0.372$) versus alphanumeric ($M = 0.334$); (2) information rate was moderate ($M = 0.385$) versus fast ($M = 0.321$); (3) three resources ($M = 0.493$) were required versus one ($M = 0.309$) or two ($M = 0.257$); (4) target shape was a triangle ($M = 0.387$) versus a circle ($M = 0.318$); and (5) target color was red ($M = 0.550$) versus blue ($M = 0.155$).

Three two-way interactions were significant. A format x color interaction resulted from the fact that more red targets were processed with the graphic ($M = 0.591$) versus alphanumeric ($M = 0.509$) representation, but for blue targets format made little difference (Ms = 0.153 and 0.158, respectively). The resources x shape interaction reflected teams' basic understanding of the targets' assigned value: for one-resource targets, teams processed a higher proportion of circles ($M = 0.382$) than triangles ($M = 0.236$); and for three-resource targets teams processed a higher proportion of triangles ($M = 0.636$) than circles ($M = 0.351$). However, even though the shape of two-resource targets did not alter their value, teams demonstrated a preference for triangles ($M = 0.291$) over circles ($M = 0.222$). A resources x color interaction indicated that the differences between the proportions of red and blue targets processed increased with the number of required resources (M differences = 0.262, 0.391, and 0.532 for one, two, and three-resource targets, respectively). The format x setting x information rate interaction is noteworthy because it was the only significant effect on targets completed which includes the group versus isolated setting variable. Inspection of Figure 3 suggests that, in the

TABLE 5. SIGNIFICANT RESULTS OF 2 x 2 x 2 x 3 x 2 x 2 WITHIN TEAM ANALYSIS OF PROPORTION OF TARGETS COMPLETED

Source	F	$p<$
Main Effects		
Format	14.55	.01
Information Rate	277.81	.01
Resources	54.29	.01
Shape	28.05	.01
Color	540.57	.01
Two-Way Interactions		
Format x Color	25.66	.01
Resources x Shape	93.65	.01
Resources x Color	134.70	.01
Three-Way Interactions		
Format x Setting x Information Rate	4.67	.05
Format x Information Rate x Color	7.58	.05
Information Rate x Resources x Shape	7.33	.01
Information Rate x Resources x Color	3.59	.05
Resources x Shape x Color	15.15	.01
Four-Way Interactions		
Information Rate x Resources x Shape x Color	7.03	.01

Note: All F tests used the mean squares of the higher order interaction with team as the error term.

worst case combination of fast information rate with alphanumeric format, the group setting facilitiated target completion (M = 0.308) relative to the isolated setting (M = 0.286).

A format x information rate x color interaction indicated that the previously reported finding, that graphic format increases the proportion of red

targets completed, was stronger for the fast information rate (M difference = 0.104) than for the moderate Information Rate (M difference =0.060).

Of the remaining significant findings, three three-way interactions and a significant four-way interaction, none included format or setting, but all included resources. Therefore, to help clarify the nature of these higher order interactions separate simple effects analyses for information rate, shape, and color at each level of resources were conducted. The cell means associated with these analyses are presented in Table 6.

TABLE 6. PROPORTION OF TARGETS COMPLETED AS A FUNCTION OF INFORMATION RATE x RESOURCES x SHAPE x COLOR

Information Rate	Color	Fast Circles	Fast Triangles	Moderate Circles	Moderate Triangles
One-Resource Targets	Blue	.183	.115	.264	.149
	Red	.502	.307	.579	.370
Two-Resource Targets	Blue	.027	.046	.079	.092
	Red	.357	.449	.427	.576
Three-Resource Targets	Blue	.047	.340	.066	.455
	Red	.631	.844	.658	.904
		Alphanumeric Group	Alphanumeric Isolated	Graphic Group	Graphic Isolated
Information Rate	Moderate	.368	.372	.404	.396
	Fast	.308	.286	.343	.345

Similar to the main effects previously reported, the main effects for information rate (more targets completed at the moderate rate) and color (more red targets completed) were significant within each level of resources. Also in accordance with the previously reported resources x shape interaction, a significant main effect for shape occurs at each level of resources, but the nature of the effect, as previously described, depends on the number of required resources (more circles completed with one resource and more triangles completed with two or three resources). Of primary concern in the present analyses are the potential interactive effects of information rate, color, and Shape at each level of resources.

For one-resource targets, only the color x shape interaction was significant ($p < 0.01$); while for blue targets the effect for shape was relatively modest (Ms = 0.224 and 0.132 for circles and triangles, respectively), for red targets the effect of shape was more extreme (Ms = 0.540 and 0.339, respectively).

The color x shape interaction was also significant for two-resource target ($p < 0.01$). Again, for blue targets the effect for shape was relatively modest (Ms = 0.053 and 0.069 for circles and triangles, respectively), while for red targets the effect of shape was more extreme (Ms = 0.392 and 0.512, respectively). The only other significant interaction for two-resource targets was color x information rate ($p < 0.05$). Analogous to the previous interactions, for blue targets the effect for information rate was relatively modest (Ms = 0.036 and 0.086 for fast and moderate rates, respectively), while for red targets the effect of information rate was more extreme (Ms = 0.403 and 0.501, respectively). For three-resource targets, the only significant interaction was information rate x shape ($p < 0.01$). For circles, the effect of information rate was modest (Ms = 0.339 and 0.362 for fast and moderate rates, respectively), while for triangles the effect for information rate was more substantial (Ms = 0.403 and 0.501, respectively).

Generally, teams performing the TRAP adequately responded to the task's constraints. Given that more targets were available than could possibly be processed, teams selected targets in a reasonable manner. As expected, teams had a marked preference for the more valuable red targets over the blue targets. Teams also demonstrated a preference for three-resource targets which integrated the use of resources more easily than searching for appropriate combinations of one-resource and two-resource targets. Finally, teams were sensitive to the targets' shape and the fact that one-resource circles were more valuable than triangles, but three-resource triangles were more valuable than circles.

Interestingly, although target shape had no bearing on the value of two-resource targets, teams nevertheless processed the red two-resource triangles more frequently than the red two-resource circles. This may be due to

the importance team members placed on processing the highly valued three-resource red triangles. That is, while looking for three-resource red triangles teams may have been more likely to notice and choose the two-person red triangles than the two-resource red circles.

Similarly, teams demonstrated biased preferences for (a) one-resource red triangles over equally valuable one-resource blue circles; and (b) three-resource red circles over equally valuable three-resource blue triangles. Therefore, teams demonstrated an overreliance on the simple color cue to the exclusion of the more complex analysis based on targets' shape and required resources.

Many of the obtained higher order interactions suggested that the effects of the independent variables tended to be multiplicative rather than merely additive. For example, differences due to format and resources were more robust for red versus blue targets. Also, for both one- and two-resource targets, the effects for shape were stronger for red versus blue targets. Similarly, effects for target variables (i.e., shape and color) tended to be enhanced with a moderate, as opposed to a fast, information rate. However, it was only under the worst-case combination of alphanumeric representation at a fast information rate that the primary independent variable, team setting, seemed to affect performance. Under such difficult circumstances a group versus isolated setting may be advantageous.

POSTEXPERIMENTAL QUESTIONNAIRE RATINGS

A postexperimental questionnaire (see Appendix C) was administered to subjects following their final test session to obtain subjective assessments of the display representations (alphanumeric versus graphic) and operator/display settings (shared versus isolated). Consistent with the performance and workload data, the graphic representation was rated more highly than the alphanumeric representation. Using a 1 to 7 scale, the graphic versus alphanumeric representation was found by subjects to be more comfortable (means = 5.7 versus 4.6), to afford greater ease of communication (means

= 5.2 versus 4.5), to supply information more adequately (means = 5.0 versus 4.8) and legibly (means = 6.0 versus 4.9), to better support both team coordination (means = 5.6 versus 4.4) and individual problem solving (means = 5.2 versus 4.5), to better facilitate understanding of the TRAP (means = 5.5 versus 4.6) and to make it an easier task (means = 5.0 versus 4.4). For all of the above comparisons, $p < 0.01$. Clearly, subjects had a strong and wide ranging preference for the graphic representation.

The only rating significantly affected by operator/display setting concerned how the communication system (headphones, microphones) affected team coordination. Subjects believed that the communication system facilitated team coordination more for the isolated (mean = 6.0) versus shared (mean = 5.3) setting ($p < 0.01$), perhaps since only in the isolated setting was the system needed by subjects to communicate with one another.

Finally, also consistent with the performance data, subjects found the task easier for trials with a moderate (mean = 5.3) versus fast (mean = 4.1) information rate ($p < 0.01$). There were no other significant effects for questionnaire ratings.

Section 8
CONCLUSIONS

Performance scores and workload ratings indicate that teams working with the graphic format or moderate information rate not only were able to perform the task more effectively, but also experienced lower workload than when working with the alphanumeric format or fast information rate. An increase of 10 in performance score corresponded with a decrease of 3.9 in SWAT rating.

Interactions in the performance data further indicate that the detrimental effect of alphanumeric format became even greater when this format was combined with the fast information rate and the individual operator/display setting.

The results of this study suggest that either multioperator/display setting may be appropriate for normal conditions involving low to moderate stress. This allows flexibility to design for other constraints such as space availability, proximity of various team members, and available display equipment. However, for a crisis condition, an organizational design which allows team members to work together in a common setting may yield better decision making performance.

REFERENCES

Baumeister, R. F. and Hutton, D. G., 1987, Self-Presentation Theory: Self-Construction and Audience Pleasing. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

Borg, G., 1978, "Subjective Aspects of Physical and Mental Load," Ergonomics, 21, 215-220.

Brown, C. E., and Leupp, D. G. (1985). Team performance with large and small screen displays. Armstrong Aerospace Medical Research Laboratory technical report number TR-85-033. (Available from NTIS.)

Carlsmith, J. M., Ellsworth, P. C., and Aronson, E., 1976, Methods of Research in Social Psychology, Reading, Massachusetts: Addison-Wesley.

Coombs, C. H., Dawes, R. M., and Tversky, A., 1971, Mathematical Psychology: An Elementary Introduction, Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Daryanian, B., February 1980, Subjective Scaling of Mental Workload in a Multi-Task Environment, Unpublished Master's Thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology.

DeSanctis, G., 1984, Computer Graphics as Decision Aids: Directions for Research, Decision Sciences, 15(4), 463-487.

Dickson, Gary W., DeSanctis, G., and McBride, D. J., 1985, Understanding the Effectiveness of Computer Graphics for Decision Support: A Cumulative Experimental Approach, Department of Management Sciences, University of Minnesota.

Dyer, J. L., 1984, Team Research and Team Training: A State-of-the-Art Review, Chapter 8 in F. Muckler (Ed.), Human Factors Review, Santa Monica, California: The Human Factors Society.

Eggemeier, F. T., 1981, Current Issues in Subjective Assessment of Workload, Proceedings of the Human Factors Society 25th Annual Meeting.

Eggemeier, F. T., Crabtree, M, S., Zingg, J. J., Reid, G. B., and Shingledecker, C. A., 1982, Subjective Workload Assessment in a Memory Update Task, Proceedings of the Human Factors Society 27th Annual Meeting, Santa Monica, California: Human Factors Society, 139-143.

Eggemeier, F. T., McGhee, J. Z., and Reid, G. B., May 1983, The Effects of Variations in Task Loading on Subjective Workload Rating Scales, Proceedings of the National Aerospace and Electronics Conference.

Eggemeier, F. T. and Stadler, M. A., October 1984, Subjective Workload Assessment in a Spatial Memory Task, Proceedings of the Human Factors Society 28th Annual Meeting.

Fiske, S. T. and Taylor, S. E., 1984, Social Cognition, New York: Random House.

Gartner, W. B. and Murphy, M. R., November 1976, Pilot Workload and Fatigue: A Critical Survey of Concepts and Assessment Techniques, National Aeronautics and Space Administration, Report No. NASA-TN-D-8365.

Geen, R. G. and Bushman, B. J., 1987, Drive Theory: Effects of Socially Engendered Arousal. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

Geothals, G. R. and Darley, J. M., 1987, Social Comparison Theory: Self-Evaluation and Group Life. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

Goldin, S. E. and Thorndyke, P. W., 1980, An Overview of Workshop Issues. In S. E. Goldin and P. W. Thorndyke (Eds.), Improving Team Performance, Proceedings of the Rand Team Performance Workshop, R-2606-ONR, Santa Monica, California: Rand.

Hackman, J. B. and Morris, C. G., 1975, Group Tasks, Group Interaction Process, and Group Effectiveness: A Review and Proposed Integration. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 8), New York: Academic Press.

Hicks, T. G. and Wierwille, W. W., 1979, Comparison of Five Mental Workload Assessment Procedures in a Moving-Base Driving Simulator, Human Factors, 21(2), pages 129-143.

Hill, G. W., 1982, Group Versus Individual Performance: Are N + 1 Heads Better Than One? Psychological Bulletin, 91, 517-539.

Jackson, J. M., 1987, Social Impact Theory: A Social Forces Model of Influence. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

Johannsen, G., Moray, N., Pew, R., Rasmussen, J., Sanders, A., and Wickens, C., 1979, Final Report of the Experimental Psychology Group, In N. Moray (Ed.), Mental Workload: Its Theory and Measurement, New York: Plenum Press.

Kahneman, D., Slovic, P., and Tversky, A., 1982, In D. Kahneman, P. Slovic, and A. Tversky (Eds.), Judgment Under Uncertainty: Heuristics and Biases, MA: Cambridge University Press.

Krantz, D. H. and Tversky, A., 1971 Conjoint Measurement Analysis of Composition Rules in Psychology, Psychological Review, 78, pages 151-169.

Lanzetta, J. T. and Roby, T. B., 1956, Effects of Work-Group Structure and Certain Task Variables on group performance, Journal of Abnormal and Social Psychology, 53, 307-314.

Lanzetta, J. T. and Roby, T. B., 1957, Group Learning and Communication as a Function of Task and Structure Demands, Journal of Abnormal and Social Psychology, 55, 121-131.

Leavitte, H. J., 1951, Some Effects of Certain Communication Patterns on Group Performance, Journal of Abnormal and Social Psychology, 46, 38-50.

McGrath, J. E., 1984, Groups: Interaction and Performance, Englewood Cliffs, New Jersey: Prentice-Hall.

Moray, N., May 1980, Subjective Measurement of Mental Workload, Arlington, Virginia: Office of Naval Research, Contract N00014-77-C-0256.

Moray, N., 1982, Subjective Mental Workload, Human Factors, 24, 25-40.

Morrissette, J. O., 1966, Group Performance as a Function of Task Difficulty, Size and Structure of Group, Journal of Personality and Social Psychology, 3, 357-359.

Mullen, B., 1987, Self-Attention Theory: The Effects of Group Composition on the Individual. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

O'Donnell, R. D., 1979, Contributions of Psychophysiological Techniques to Aircraft Design and Other Operational Problems, NATO Advisory Group for Aerospace Research and Development, Report No. 244.

Pattipati, K. R., Kleinman, D. L., and Ephrath, A. R., 1982, A Dynamic Decision Model of Human Task Performance. University of Connecticut Technical Report, Dept. Of EEEC, TR-82-1.

Pruitt, D. G. and Kimmel, M. J., 1977, Twenty years of Experimental Gaming: Critique, Synthesis, and Suggestions for Future. Annual Review of Psychology, 28, 363-392.

Pryor, J. B. and Ostrom, T. M., 1987, Social Cognition: Theory of Group Processes. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

Reid, G. B., May 1982, Subjective Workload Assessment: A Conjoint Scaling Approach, Proceedings of the 1982 Annual Scientific Meeting of the Aerospace Medical Association.

Reid, G. B., Eggemeier, F. T., and Nygren, T. E., 1982, An Individual Differences Approach to SWAT Scale Development, Proceedings of the Human Factors Society 26th Annual Meeting. Santa Monica, California: Human Factors Society, 639-642.

Reid, G. B., Eggemeier, F. T., and Shingledecker, C. A., May 1982, Subjective Workload Assessment Technique, Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics.

Reid, G. B., Shingledecker, C. A., and Eggemeier, F. T., October 1981, Application of Conjoint Measurement to Workload Scale Development, Proceedings of the Human Factors Society 25th Annual Meeting.

Reid, G. B., Shingledecker, C. A., Hockenberger, R. L., and Quinn, T. J., May 1984, A Projective Application of the Subjective Workload Assessment Technique, Proceedings of the IEEE 1984 National Aerospace and Electronics Conference.

Reid, G. B., Shingledecker, C. A., Nygren, T. E., and Eggemeier, F. T., 1981, Development of Multidimensional Subjective Measures of Workload, Proceedings of the 1981 IEEE International Conference on Cybernetics and Society, 403-406.

Sande, G. N. and Zanna, M. P. (1987). Cognitive Dissonance Theory: Collective Actions and Individual Reactions. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

Sanders, A. F., 1979, Some Remarks on Mental Load, In N. Moray (Ed.), Mental Workload: Its Theory and Measurement, Plenum Press: New York.

Schwartz, David R. and Howell, William C., 1985, Optional Stopping Performance under Graphic and Numeric CRT Formatting, Human Factors, 27(4). 433-444.

Scott, Brad and Wickens, Christopher D., 1983, Spatial and Verbal Displays in a C^3 Probablistic Information Integration Task, Proceedings of the Human Factors Society 27th Annual Meeting, 355-358.

Shaw, M. E., 1976, An Overview of Small Group Behavior, In J. W. Thibaut, J. T. Spence, and R. C. Carson (Eds.), Contemporary Topics in Social Psychology, Morristown, New Jersey: General Learning Press.

Sheridan, T. B. and Simpson, R. W., January 1979, Toward the Definition and Measurement of Mental Workload of Transport Pilots, FTL Report R79-4, Massachusetts Institute of Technology Flight Transportation Laboratory: Cambridge.

Shiflett, S., 1979, Toward a Generral Model of Small Group Productivity, Psychological Bulletin, 86, 67-79.

Shingledecker, C. A., Crabtree, M. S., and Acton, W. H., 1982 Standardization Tests for the Evaluation and Classification of Workload Metrics, Proceedings of the Human Factors Society 26th Annual Meeting.

Stasser, G. and Titus, W., 1985, Pooling of Unshared Information in Group Decision Making: Biased Information Sampling During Discussion, Journal of Personality and Social Psychology, 48, 1467-1478.

Steiner, I. D., 1972, Group Process and Productivity, New York: Academic Press.

Stollings, M. N., 1982, Information Processing Load of Graphics Versus Alphanumeric Weapon Format Displays for Advanced Fighter Cockpits, Master's Thesis. Wright State University, Dayton, Ohio.

Wegner, D. M., 1987, Transactive Memory: A Contemporary Analysis of the Group Mind. In B. Mullen and G. R. Geothals (Eds.), Theories of Group Behavior, New York: Springer-Verlag.

Wierwille, W. W., October 1979, Physiological Measures of Aircrew Mental Workload, Human Factors, 21, 575-594.

Wierwille, W. W. and Connor, S. A., 1983, Evaluation of 20 Workload Measures Using a Psychomotor Task in a Moving-Base Aircraft Simulator. Human Factors, 25, 1-16.

Williges, T. C and Wierwille, W. W., 1979, Behavioral Measures of Aircrew Mental Workload, Human Factors, 21, 549-574.

Appendix A

CONSENT FORM

I, ____________________, having full capacity to consent, do hereby volunteer to participate in a research study entitled, "Team Resource Allocation Problem," under the direction of Ms. Denise L. Wilson, Mr. Michael McNeese, Dr. Clifford Brown, and Lt. Suzanne Kelly. The decision to participate in this research is completely voluntary on my part. No one has coerced or intimidated me into participating in this program. I am participating because I want to. ____________________ has adequately answered any and all questions I have about this study, my participation, and procedures involved. I understand that ____________________ will be available to answer any questions concerning procedures throughout this study. I understand that if significant new findings develop during the course of this research which may relate to my decision to continue participation, I will be informed. I further understand that I may withdraw this consent at any time and discontinue further participation in this study without prejudice to my entitlements. I also understand that the medical monitor of this study may terminate my participation in this study if he or she feels this to be in my best interest.

I understand that my participation in this study may be photographed, filmed, or videotaped. I consent to the use of these media and understand that any records of my participation in this study may only be disclosed according to federal law, including the Federal Privacy Act, 5 USC 552a, and its implementing regulations.

I understand that my entitlement to medical care or compensation in the event of injury is governed by federal laws and regulations, and if I desire further information I may contact ____________________.

I FULLY UNDERSTAND THAT I AM MAKING A DECISION WHETHER OR NOT TO PARTICIPATE. MY SIGNATURE INDICATES THAT I HAVE DECIDED TO PARTICIPATE HAVING READ THE INFORMATION PROVIDED ABOVE.

AM
PM

Signature Date Time

I was present during the explanation referred to above, as well as the volunteer's opportunity for questions, and hereby witness the signature.

Signature Date

I have briefed the volunteer and answered questions concerning the research project.

Signature Date

ADDENDUM TO THE CONSENT FORM

Experiment: Team Resource Allocation Problem

You are invited to participate in an experiment designed to study how people in teams work with one another to complete a number of tasks. The situation you will be exposed to has theoretical similarity to those encountered in command, control, and communications (C^3) systems of the USAF. A better understanding of the processes by which team members complete tasks will assist in improving these C^3 systems. Your exposure to the equipment is limited to your watching the CRT screen at a distance of about 2 feet for approximately 1 hour per day for 4 days. This does not involve any known risks.

In the experiment, you will be observing a computer generated display of a representation of a work environment. By pressing pushbuttons on a response box, you will work on tasks individually and with your team members. Because there will be more tasks available to you than you can complete, the particular tasks you and your team members choose, and when you choose them, will be of primary interest. You will receive further detailed instructions at the beginning of the experiment.

The responses you make, and the times at which you make them, will be recorded for later analysis. Audio and video recordings will also be made for subsequent study. Your name will be recorded along with the dates and times at which the experiment is performed. Your confidentiality as a participant in this project will be protected. Your identity will only be revealed in accordance with the Privacy Act, 5 USC 552, and its implementing regulations. A numeric code will be used to identify the data in any publication.

Any monetary benefits will be in accordance with SRL/Air Force agreements.

You are free to refuse to participate or to withdraw your participation in the experiment at any time. Doing so will not prejudice your relation with the Laboratory in any respect.

Any questions you may have should be directed to Ms. Denise Wilson (57572) or Mr. Michael McNeese (58805).

Your willingness to participate in this experiment is greatly appreciated. Your signature indicates that you have decided to participate, having read the information provided above.

YOU WILL BE GIVEN A COPY OF THIS FORM TO KEEP.

Date Volunteer's Initials

Appendix B
INSTRUCTIONS

TEAM RESOURCE ALLOCATION PROBLEM

Before you take seats, please take a card which will determine whether you will be team member A, B, or C for the study. (Subjects draw cards and are seated at appropriate seats.) (Explain seating/display arrangement.)

LARGE SCREEN SHARED DISPLAY

You and your teammates will share the single large screen display. You will communicate with your teammates through the headphones provided at the workstation. Each team member has an individual control box at his/her workstation.

INDIVIDUAL CRT DISPLAYS

Each team member has his/her own display and control box at the individual workstations. You will communicate with your teammates through the headphones provided at the workstation.

GENERAL INSTRUCTIONS

This experiment is concerned with how team members work with one another to accomplish tasks. You and your teammates will work together to decide how best to allocate team resources (your work time) for the good of the team in a task which involves the processing of various targets. You will work on some targets yourself, and on other targets with one or both of your teammates. Over the course of the experiment, you will see various representations of your team's work environment. However, the basic task will always be the same. The major portion of the display will have 11 rows. Targets, represented as RED and BLUE CIRCLES and TRIANGLES appear randomly in each of these 11 rows. You will earn points for the team by working on these targets before their time runs out and they leave the screen.

Working on a target simply means selecting a target by using the buttons on your control box to move a cursor, pressing the start button, and waiting a few seconds for the target to be processed. Because more targets than you can possibly work on will appear on the screen, the particular targets you choose, and the point value to the team, will be quite important. Therefore, it is necessary for you to learn how the point values of targets are determined. Please listen carefully.

The point value of each target depends on three things: the number of required workers, the color of the target, and the shape of the target. Overall, the point value of a target is proportional to the number of required workers. The average value of all the targets is three points per person. Targets requiring one, two, or three workers are, therefore, worth an average value of three, six, or nine points, respectively, for the team. Whether a particular target is worth this average value of three points per person, or somewhat more or less, depends on its color and shape.

Since the color RED suggests importance or urgency and the color BLUE suggests calm, you will see that RED targets are worth more points on the average than BLUE targets. Since a circle suggests unity or oneness, and a TRIANGLE, having three corners, suggests the number three, you will see a circle when one worker is required. CIRCLES are worth more points than TRIANGLES; but when three workers are required, TRIANGLES are worth more points than CIRCLES. With these notions of color and shape in mind, let us examine the specific point values assigned to the different targets. (Give subjects the point values table.)

Overall, RED targets are worth four points per person, and BLUE targets are worth two points per person. One way to remember this is that there is a one point per person bonus for completing RED targets, and a one point per person penalty for completing BLUE targets. Therefore, RED targets are worth the average value (three points per person) plus the bonus (one point per person) which equals four points per person. Similarly, BLUE targets are worth the average value (three points per person) minus the penalty (one point per person) which equals two points per person.

However, the shape of the target also influences the target's point value. For one-person targets, there is a one point bonus for CIRCLES and a one point penalty for TRIANGLES. For BLUE one-person targets (which are worth two points per person on the average), a BLUE TRIANGLE is worth one point, but a BLUE CIRCLE is worth three points. Similarly, for RED one-person targets (which are worth four points per person on the average), a RED TRIANGLE is worth three points, but a RED CIRCLE is worth five points. For a one-person target, a BLUE CIRCLE and a RED TRIANGLE are worth the same, three points. This is because each has both a bonus and a penalty which cancel each other out.

When a target is processed by two persons, the shape of the target does not matter. RED CIRCLES and TRIANGLES are worth eight points (four points per person), and BLUE CIRCLES and TRIANGLES are worth four points (two points per person). Therefore, when targets require two workers, only the color of the target determines its points value.

For three-person targets, there is a one point per person bonus for TRIANGLES and a one point per person penalty for CIRCLES. For BLUE three-person targets (which are worth two points per person on the average), a BLUE CIRCLE is worth three points (one point per person) but a BLUE TRIANGLE is worth nine points (three points per person). Similarly, for RED three-person targets (which are worth four points per person on the average), a RED CIRCLE is worth eight points (three points per person), but a RED TRIANGLE is worth 15 points (five points per person). For a three-person target, a BLUE TRIANGLE and a RED CIRCLE are worth the same, nine points (three points per person). This is because each has both a bonus and a penalty which cancel each other out.

You should now be able to determine the point value of each target by knowing the number of required workers and the target's color and shape. Let me summarize what you have learned. You begin with the idea that each target is worth three points per person. To this, you add one point per person if it is RED, or subtract one point per person if it is BLUE. Finally, you may have to apply a one point per person bonus or penalty based on the target's shape. For one-person targets, CIRCLES have a

one-point per person bonus and TRIANGLE a one-point per person penalty. For three-person targets, TRIANGLES have a one point per person bonus and CIRCLES have a one point per person penalty. For two-person targets, only the color determines point value, shape does not matter.

Because your understanding of the point values is critical to this study, I am going to have you complete a short test to demonstrate your knowledge of the point values of each target. Before taking the test please examine the summary table of the point values and feel free to ask questions about it. (Pause) Do you have any questions before you take the test?

(Give subjects the test. If any questions are missed, discuss the question with the subject to ensure his understanding and then give him a new test. Repeat this procedure until all subjects have answered all the questions correctly.)

We are ready to continue. (START APPUPRIATE DEMO)

GRAPHIC FORMAT

The TRAP task is represented here in its graphic format. As you can see, there are 11 rows on which targets, RED and BLUE CIRCLES and TRIANGLES appear at random and move across the screen from left to right. The black squares in columns A, B, and C indicate which operators are required to work on each target. The scale at the top represents 30 time units.

Working on a target is very simple. All you do is move your marker, a green asterisk, to a target row and press the start button on your response box. Work automatically begins, and after a short time (10 time units), a beeping noise will indicate that your team received the appropriate number of points for completing the target. These points are automatically added to the accumulated points display (show). When you begin work on a target, a black rectangle will appear in that target's row. The rectangle represents the 10 time units required to complete processing of the target. The target will move through the rectangle as the processing proceeds. When the target moves out of the rectangle, processing of the target has been

completed. In order to complete a target before it leaves the screen, you will have to start it before it reaches the black dashed line (while the target is in the opportunity window).

Work on each target can be done only by a particular team member or combination of team members. As the control box before you indicates, you are either team member A, B, or C. You can work only on those targets which have a black square in your column. If a target has more than one black square in front of it, both or all three corresponding team members will have to work on the target <u>at the same time</u> in order to complete it.

To work on a target, you must move your marker to the corresponding black square. You move the marker by pressing the buttons labeled up and down on your control box. Go ahead and move your marker around. Notice that if you press the up button when you are on the top row, your marker moves to the bottom row. Similarly, if you press the down button when you are on the bottom row, the marker moves to the top row.

Once you have the marker on the row correspnding to the target you wish to work on, all you have to do is press the start button. If you are the only team member required for that target, work automatically begins and the black square will turn yellow. However, if one or more additional team members are required for the target, the black square will turn pink. This means that you are waiting to work. Work will begin only when all the required workers for the target have moved their markers to the target and pressed their start buttons. When this occurs, all the squares will turn yellow indicating that work has begun.

While you are working on a target or waiting for another team member at a target, your marker will turn red. You can move it to any row you choose in preparation for the next target you may wish to start. When you become free, your marker will return to its green color indicating that you are ready to press the start button for another target.

You may wish to stop working on a target before completing it. To do this, you simply press the RESET button on your control box. Your marker will

turn green indicating that you are free to start another target. If others were working on the target with you, they will also have to press their reset buttons to work on a different target. You will receive no points for targets which are not fully completed. If you choose, you may begin to process the target over again, but it will take a full 10 time units to complete it. The reset button is also used when you no longer wish to wait for other team members at a particular target.

Processing of each target takes 10 time units (TUs). A TU is some arbitrary number of seconds. The current example trial has a TU of 3 seconds. It takes 30 TUs (in this example, 90 seconds) for a target to move completely across the screen. During the actual experiment, the number of seconds for a TU will be less. The targets will move across the screen more quickly, and the time spent processing each target will be less.

The table in the lower left hand portion of the display indicates whether each team member is free, waiting, or working. A black square indicates that a particular team member is free, while a blinking pink square indicates that a particular team member is waiting. When a particular team member is working, a numeric countdown, in TUs, will indicate how much processing time remains until the team member will be finished with the current target.

The countdown for each target starts at 30 TUs when a target is at the left-most part of the screen, and decreases at a constant rate as the target moves to the right. When a target is at the end of the opportunity window (the black dashed line), the countdown will be at 10 TUs. The target leaves the screen at 0 TU. Since each and every target requires 10 TUs for processing, knowing how many TUs a target will remain on the screen can be useful to you as you decide which targets to work on, and when to work on them. In addition, comparing this information to the countdown of team members who are currently working (show), can provide actual information about whether there will be enough time to process particular targets. For example, if a team member has 6 TUs remaining before completing a particular target, he will not be able to complete both that target and another target that currently has only 15 TUs remaining.

The object of this exercise is to accumulate as may points as possible as a team. This means discussing alternatives with the other members of the team in order to make optimum selection of targets. As there will be more targets than the team can possibly process, combinations of targets should be selected which optimize team performance and total point count.

ALPHANUMERIC FORMAT

The TRAP task is represented here in alphanumeric format. As you can see, there are 11 rows in which targets may appear at random. The black squares in columns A, B, and C indicate which operators are required for processing a particular target. The next column shows the number of points the team can earn for processing that target. The two-part target names (COLOR and SHAPE) are given in the next column followed by the time available to process that target. The last column indicates the status of the target if it is being worked or waiting for an operator.

Work on each target can be done only by a particular team member or combination of team members. As the control box before you indicates, you are either team member A, B, or C. You can work only on those targets which have a black square in your column. If a target has more than one black square in front of it, both or all three corresponding team members will have to work on the target at the same time in order to complete it.

To work on a target, you must move your marker, a white asterisk (*) to the corresponding black square. You move the marker by pressing the buttons labeled up and down on your control box. Go ahead and move your marker around. Notice that if you press the up button when you are on the top row, your marker moves to the bottom now. Similarly, if you press the down button when you are on the bottom row, the marker moves to the top row.

Once you have the marker on the row corresponding to the target you wish to work on, all you have to do is press the start button. If you are the only team member required for that target, work automatically begins and the entire row will be displayed in reverse video and the status column will display "working" for that target. However, if one or more additional team

members are required for the target, the status column will display a blinking "waiting" for this target. This means that you are waiting to work. Work will begin only when all the required workers for the target have moved their markers to the target and pressed their start buttons. When this occurs, the row will be displayed in reverse video and the status column will display "working" indicating that work has begun.

While you are working on a target or waiting for another team member at a target your marker will turn to a circle (o). You can move it to any row you choose in preparation for the next target you may wish to start. When you become free, your marker will return to its asterisk (*) form indicating that you are ready to press the start button for another target.

You may wish to stop working on a target before completing it. To do this, you simply press the RESET button on your control box. Your marker will turn to an asterisk indicating that you are free to start another target. If others were working on the target with you, they will also have to press their reset buttons to work on a different target. You will receive no points for targets which are not fully completed. If you choose, you may begin to process and target over again, but it will take a full 10 TUs to complete it. The reset button is also used when you no longer wish to wait for other team members at a particular target.

Processing of each target takes 10 TUs. A TU is some arbitrary number of seconds. The current example trial has a TU of 3 seconds. Each target is available for 30 TUs (in this example, 90 seconds). During the actual experiment, the number of seconds for a TU will be less. That is, the target's countdown will proceed more quickly, and the time spent processing each target will be less.

The table in the lower left hand portion of the display indicates whether each team member is free, waiting, or working. A black square indicates that a particular team member is free, while a blinking "W" indicates that a particular team member is waiting. When a particular team member is working, a numeric countdown, in TUs, will indicate how much processing time remains until the team member will be finished with that target.

The countdown for each target will start at 30 TUs when a target is first displayed, and decreases at a constant rate as the 30 TUs available for processing the target elaspe. When the countdown is at 10 TUs, there is just enough time left to process the target. The target is deleted from the table when 0 TUs remain. Since each and every target requires 10 TUs for processing, knowing how many TUs remain for each target can be useful to you as you decide which targets to work on, and when to work on them. In addition, comparing this information to the countdown of team members who are currently working (show) can provide vital information about whether there will be enough time to process particular targets. For example, if a team member has 6 TUs remaining before completing a particular target, he will not be able to complete both that target and another target that currently has only 15 TUs remaining.

The object of this exercise is to accumulate as many points as possible as a team. This means discussing alternatives with the other members of the team in order to make optimum selection of targets. As there will be more targets than the team can possibly process, combinations of targets should be selected which optimize team performance and total point count.

Appendix C

TEAM RESOURCE ALLOCATION PROBLEM (TRAP) II
EVALUATION QUESTIONNAIRE (Teams 3 and 11)

Name: ______________________________ Team:_____ Operator: A B C

Date:______________

The purpose of this experiment was to investigate the effects of display format (graphic versus alphanumeric) and operator/display orientation (large screen shared display versus individual CRTs) upon team performance in a dynamic problem solving task. Your subjective impressions, as recorded on the following questionnaire, will complement the performance data. These data, and that of subsequent studies, will be used to make recommendations regarding the design of command centers. Please consider your answers carefully and add your written comments as appropriate.

Please rate the legibility of the information as presented on the display.

Session 1: Alphanumeric Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7
not at all legible ... highly legible

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7
not at all legible ... highly legible

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7
not at all legible ... highly legible

Session 4: Alphanumberic Format/Group Display

1---------2--------3---------4---------5---------6---------7
not at all legible ... highly legible

COMMENTS:____

____

____

____

Please rate the adequacy (for performance of the task) of the information presented on the display.

Session 1: Alphanumeric Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

not at all adequate — highly adequate

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

not at all adequate — highly adequate

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7

not at all adequate — highly adequate

Session 4: Alphanumberic Format/Group Display

1---------2--------3---------4---------5---------6---------7

not at all adequate — highly adequate

COMMENTS:____

____

____

____

To what extent do you feel that the communication system (headphones, microphone) affected team coordination?

Session 1: Alphanumeric Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

highly disruptive — highly facilitative

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7
highly disruptive — highly facilitative

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7
highly disruptive — highly facilitative

Session 4: Alphanumeric Format/Group Display

1---------2--------3---------4---------5---------6---------7
highly disruptive — highly facilitative

COMMENTS: ____

____

____

____

Please rate the difficulty of the two time-stress conditions of the TRAP task for each of the following:

Session 1: Alphanumeric Format/Individual CRTs

Faster trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

Slower trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

Session 2: Graphic Format/Individual CRTs

Faster trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

Slower trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

Session 3: Graphic Format/Group Display

Faster trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

Slower trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

Session 4: Alphanumeric Format/Group Display

Faster trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

Slower trials:

1---------2--------3---------4---------5---------6---------7
very difficult — very easy

COMMENTS:____

Please rate each of the experimental conditions according to how comfortable you felt.

Session 1: Alphanumeric Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7
very uncomfortable — very comfortable

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7
very uncomfortable — very comfortable

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7
very uncomfortable — very comfortable

Session 4: Alphanumeric Format/Group Display

1---------2--------3---------4---------5---------6---------7
very uncomfortable — very comfortable

COMMENTS: ____

____

____

____

Please rate each of the experimental conditions for ease of communication.

Session 1: Alphanumeric Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7
very difficult communication — very easy communication

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7
very difficult communication — very easy communication

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7

very difficult communication | very easy communication

Session 4: Alphanumeric Format/Group Display

1---------2--------3---------4---------5---------6---------7

very difficult communication | very easy communication

COMMENTS: ____

Please rate each of the experimental conditions for its effects upon team coordination.

Session 1: Alphanumeric Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

inhibits coordination | enhances coordination

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

inhibits coordination | enhances coordination

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7

inhibits coordibation | enhances coordination

Session 4: Alphanumeric Format/Group Display

1---------2--------3---------4---------5---------6---------7

inhibits coordination | enhances coordination

COMMENTS:____

Please rate each of the experimental conditions for its facilitation of the team members understanding of the TRAP.

Session 1: Alphanumeric Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

Session 4: Alphanumeric Format/Group Display

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

COMMENTS:____

Please rate the extent to which you feel each of the experimental conditions was facilitative to individual problem solving.

Session 1: Alphanumberic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

Session 2: Graphic Format/Individual CRTs

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

Session 3: Graphic Format/Group Display

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

Session 4: Alphanumeric Format/Group Display

1---------2--------3---------4---------5---------6---------7

non-facilitative ... highly facilitative

COMMENTS: ____

____

____

____

Did you find the TRAP task enjoyable? _____ How would you improve it?

____

____

____

____

Can you suggest improvements in the information portrayal (i.e., what information is displayed. How it is represented. Where it is located on the display, etc.)? ______________________________

____

____

____

Can you suggest ways to improve team communication and coordination?

____

____

____

____

Did the "accumulated points" scale give adequate feedback? What additional feedback would you like? ______________________________

THANK YOU FOR PARTICIPATING!!!!

Appendix D

GROUPS IN TRAP: AN EXPLORATORY ANALYSIS OF BEHAVIORAL DYNAMICS

A. Rodney Wellens
University of Miami

and

Tracy Vogler
Fairborn High School

This paper describes a pilot research project conducted while the senior author was a Visiting Faculty Fellow at the Armstrong Aerospace Medical Research Laboratory, Human Engineering Division, Technology Development Branch. The second author was a summer high school apprentice assigned to the same laboratory. The research was sponsored by the Air Force Office of Scientific Research/AFSC, U.S. Air Force, under Contract F4920-85-C-0013.

Groups in TRAP: An Exploratory Analysis of Behavioral Dynamics

Abstract

A pilot study was conducted that assessed group verbal and nonverbal behaviors during team problem solving activities. Eight three-person teams were videotaped while they were engaged in the Team Resource Allocation Problem (TRAP). The subjects were part of a larger study designed to assess the effects of high versus low time stress, alphanumeric versus graphic display formating, and group versus isolated viewing of data displays on total team performance.

Preliminary results suggested that of the nine verbal and nonverbal behavioral categories sampled, verbal commands and gesturing were the best indicators of group functioning and responded directly to two of the three independent variables manipulated. Verbal commands occurred at at faster rate under conditions of high as opposed to low time stress and in isolated as opposed to shared display settings. Gesturing increased dramatically as a function of group versus isolated viewing of displays. Other behaviors that were less clearly related to experimental manipulations were nonetheless helpful in defining team operating climates.

Team performance, as measured by total game points accumulated, was higher for graphics than alphanumeric displays and lower for high time stress than low time stress conditions. However, team efficiency scores, measured by the rate of point accumulation, indicated higher processing efficiency under high time stress. Directed command rates from emergent leaders appeared to covary positively with team performance under high time stress conditions and negatively under low time stress. Gesturing was negatively related to team performance. A consideration of processing opportunities (display rates) balanced against team task choices (optimal selection strategies) is necessary to develop a better measure of team performance capabilities. Additional studies using more rigorous experimental control and measurement techniques are recommended.

Groups in TRAP: An Exploratory Analysis of Behavioral Dyamics

A. Rodney Wellens and Tracy Vogler

Overview of TRAP

The Team Resource Allocation Problem (TRAP) was developed by Dr. Clifford Brown to study team performance in a multiperson, multitask setting that captures many of the essential elements of a C^3 system in which decisions are made within a dynamic, time-stressed environment. The task is a computer based cooperative problem solving game that is an extension of an experimental paradigm used by Pattipati, Kleinman, and Eprath (1982).

As currently configured, three-person teams work together to accumulate as many points as possible by processing "tasks" that are presented symbolically on a CRT screen. Tasks are processed by team members pressing buttons controlling separate cursors that appear on the CRT. Group coordination is encouraged by requiring integration of team members activities to optimize point accumulation. A more detailed explanation of the task may be found in Brown and Leupp (1985) and Section 3 of this report.

Two variations on the display of information for team members has recently been devised. An "alphanumeric" display shows a tabular arrangement of time remaining on processed tasks and opportunity windows for nonprocessed tasks by use of a series of digital countdown counters. A "graphic" version of the display shows task symbols moving horizontally across the display with an opportunity window bounded by a vertical cut-off line. Processing time remaining on chosen tasks is indicated by a darkened horizontal bar placed behind the chosen symbol. Figure D-1 shows these two variations in display format.

When the senior author arrived at AAMRL, an experiment was in progress examining the effects of high versus low time-stress, group versus isolated viewing of CRTs, and graphic versus alphanumeric display of task variables within the TRAP setting. The effort was part of a long term study of large group displays and team performance (McNeese and Brown, 1986). The

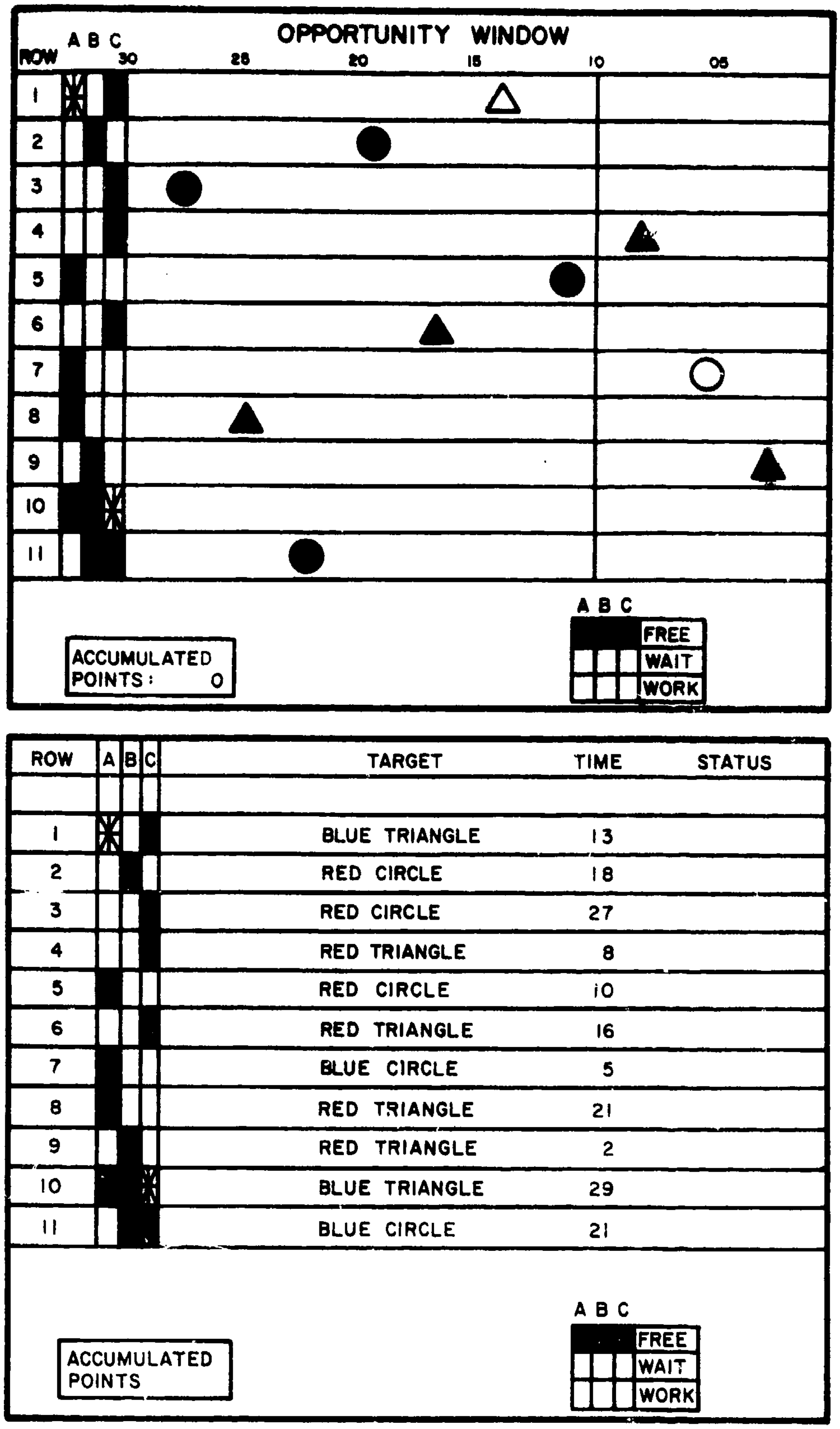


Figure D-1. Comparison of Graphic (Top) and Alphanumeric (Bottom) TRAP Display Formats

experiment used a 2 x 2 x 2 modified latin square design with repeated measures on each primary factor. Sixteen teams had been scheduled and seven teams had already been run or were midway through their experimental sessions.

Each team met on four separate occasions. On two occasions, individuals were positioned in a group orientation (seated side by side facing a large screen rear projection display). On one of these occasions, they viewed a graphic version of TRAP; on the other occasion, they viewed an alphanumeric version of TRAP. On the remaining two occasions, team members were visually isolated from one another and viewed individual CRTs, again being exposed to either a graphic or alphanumeric version of TRAP. In all conditions, team members verbally communicated with each other via head mounted microphones and earphones.

Within each session, teams were exposed to eight experimental trials, four of these trials proceeded at a slow information display rate (low time stress = 60 tasks in 328 seconds) and four at a fast display rate (high time stress = 60 tasks in 164 seconds). The main dependent variables were group performance scores (total number of points accumulated within a trial) and subjective estimates of workload (SWAT ratings; see Reid, 1982). Subjects were paid volunteers recruited from a local college population without regard to sex or prior acquaintance. All subjects displayed 20/20 visual acuity.

Exploratory Behavioral Analysis

In order to obtain relevant behavioral data regarding team coordination and communication, the senior author recommended that videotapes be made of the experimental sessions for behavioral analyses. Video recording requipment was quickly located and placed into service[1] to collect data on the

[1]The authors wish to thank MSgt. Danny Bridges for locating and assembling the necessary video recording equipment and Mr. Bill McGovern for operating the equipment during the experimental sessions. The authors would also like to thank Dr. Clifford Brown, Mr. Michael McNeese, and Mr. Donald Monk for their helpful comments on an earlier draft of this report.

remaining teams within the physical restrictions of the laboratory being used. The following represents a description of the recording and rating procedures used together with a set of impressions based upon an initial "first pass" analysis of the videotapes. Recommendations are then presented for future experiments.

Physical Setting and Recording Procedures

All TRAP sessions were conducted within an 8.5-foot x 26-foot room equipped with 6-foot 8-inch room dividers for creating the isolated viewing conditions and a General Electric 4PJ5150B5 large screen color video projection system for the group viewing conditions. The room layout is depicted in Figure D-2. During group viewing conditions, persons 'A', 'B', and 'C' were seated at tables A, B, and C that were placed approximately 8 feet in front of a rear projection screen that displayed a 4-foot x 3-foot video image. During isolated viewing conditions, persons 'A', 'B', and 'C' were seated at tables A', B', and C'. These tables held Conrac 7211C13 high resolution color monitors. During all sessions, a Panasonic 3150 color video camera mounted atop a partition along the right hand wall, approximately 4 feet from the front of the room, recorded all subject behaviors within its field of view. The camera, placed 6 feet 8 inches above the floor, was equipped with a Cosmicar 8.5 mm 1:1.5 television lens. The distance from the lens to the nose of person 'B' in the group condition was approximately 8 feet. The camera was connected to a Panasonic PV1730 VHS video cassette recorder. A Datavision DT-1 time/date generator was used to superimpose a running 0.1-second digital clock upon the top protion of all recorded video images. The head, upper torso, lap, arms, and hands of persons 'A', 'B', and 'C' could be seen in the group condition. While persons 'B' and 'C' could be seen in the isolated condition, person 'A' could not be seen at all.

In all conditions, team members wore Astrocom 20680 headsets equipped with matching Electro-voice dynamic microphones. A Tascam 44 four-channel tape recorder was used to mix audio signals and interconnect the participants. The mixed audio was fed to the Panasonic VCR for recording on one audio channel.

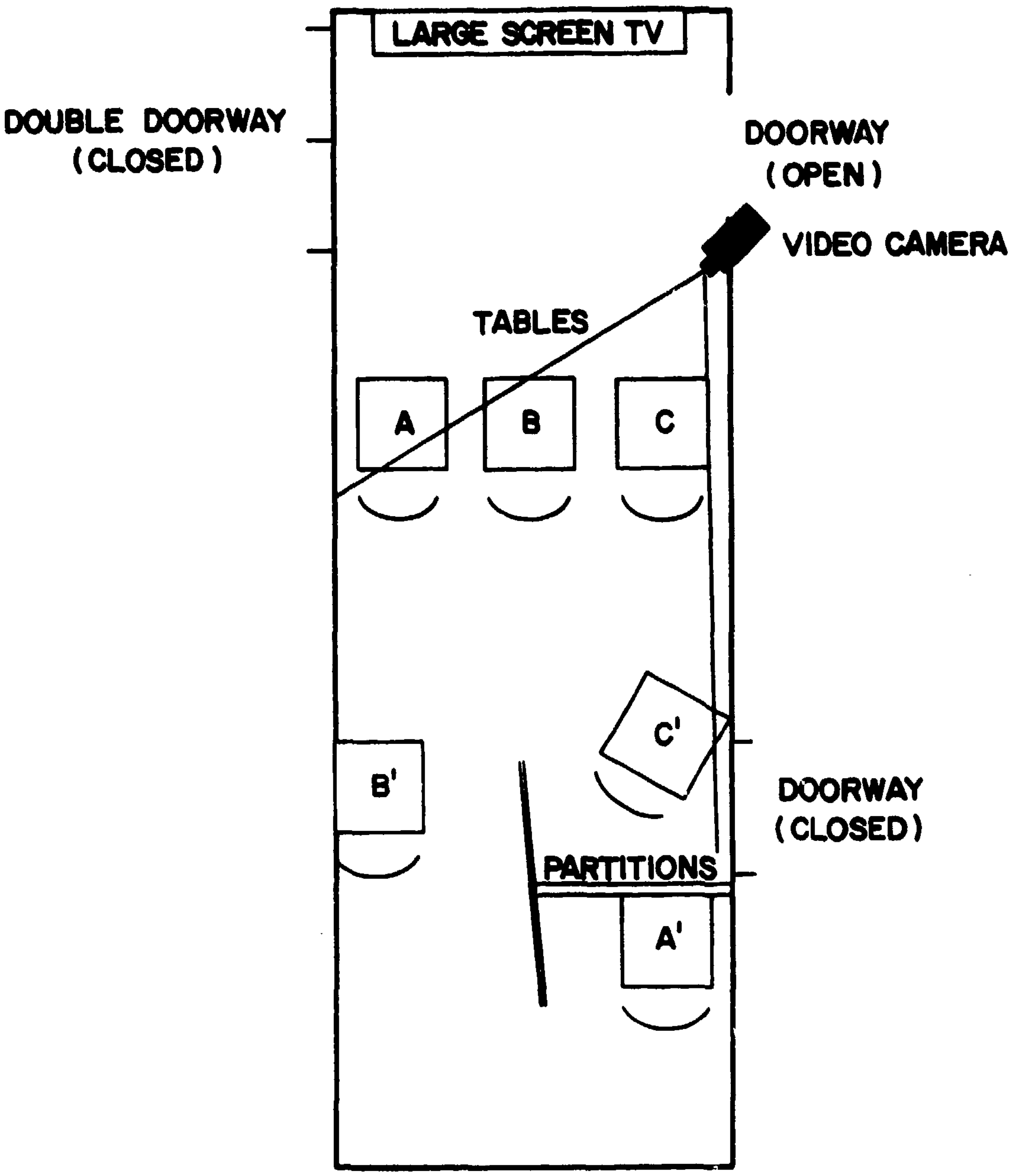


Figure D-2. Floor Diagram of Experimental Setting Showing Camera Placement and Group (A, B, and C) Versus Isolated (A', B', and C' Seating Arrangements (Lines from Camera Show Field of View)

Observational Rating Procedure

Because videotaping began at the approximate midpoint of the ongoing experiment, the first videotapes that became available for observation were made of groups who were already in the midst of their second to fourth experimental session. These initial tapes were not included in the final set to be analyzed but were reviewed to determine the range of behaviors to be expected from new teams that were being scheduled. Because of the limited amount of time available for rating, an abbreviated rating system that drew upon Bales' (1950) interaction process analysis technique was developed that simply counted the number of occurrences of certain verbal and nonverbal behaviors that could be taken as indicators of team coordination efforts (e.g., giving verbal commands, asking questions, pointing at information on the large screen CRT) and affective expressions/or tension release (positive or negative interpersonal comments, self manipulations, posture changes, yawning, laughing). The behavioral taxonomy was by no means exhaustive.

A copy of the rating form developed is included in Attachment A. The form was divided according to rating categories and subdivided into rating units (each small block representing one occurrence of a behavior). Again, because of time restraints, only the first four trials of each experimental session was rated. This included two trials of high and low time stress per session as well as the SWAT rating time between trials. Because the rater could not see the visual display attended to by team members, he had to estimate the beginning and ending points of trials by observing clock times and by observing team members behaviors. Future studies should include either a superimposed view of the display to give raters ready access to TRAP information or at least provide an auditory cue for marking stop and start points.

After a brief training session, the second author proceeded to view all videotapes that included four experimental sessions per team. Earphones were attached to the VCR to aid listening to verbal comments made by team members. Each time the rater detected a verbal or nonverbal behavior that fit into the rating scheme, the video tape was freeze-framed and the

behavior tallied on the rating sheet. A total of eight teams were observed through the first half of each of their four 1-hour sessions. Because of the "stop and go" nature of the rating procedure, the first pass through the video tapes consumed approximately 60 hours of observer time. Limited time resources prohibited a second pass by an independent rater. Therefore, no estimate of interrater reliability was calculated. However, the single rater did review several sessions to gauge his own consistency and reported confidence in the relative, if not absolute, frequencies of the most frequently recorded behaviors: verbal commands, self manipulations, and gesturing. The rater felt least confident about the verbal category of questions asked, positive and negative comments, and the nonverbal categories of smiling/laughing and yawning. Obviously, a second pass through the videotapes will be necessary to verify the tentative results reported here.

Preliminary Results

Raw Frequency Data

An example of the rater's raw tally sheets appears in Attachment B. The rating sheets were specifically designed to be read as histograms depicting the relative frequency of target behaviors for individuals within teams for each trial of each experimental session. A review of the raw data sheets revealed large individual differences in the frequency of verbal commands elicited by subjects. As trials progressed for most teams, a single dominant "leader" emerged that tended to direct the other team members' activities. This was evidenced by a marked difference between this team member's verbal activity level compared to his/her temamates. This leader persisted across sessions and experimental conditions.

Examination of the raw data sheets also revealed that the intertrial "break" time (when subjects were supposed to be making their SWAT ratings) was actually quite active. Positive or negative comments about performance scores were most often exchanged during this period. Some sharing of subjective workload estimates was also observed during this time.

Occasionally, team members would point out where other team members should look for information on the large screen display during "break" time. Posture changes, accompanied by yawning and stretching tended to occur at the end of trials and during break periods. Self manipulations (i.e., rubbing the face, hands, or legs) occurred both during break periods and during actual trials. In general, subjects tended to focus intently upon their visual displays during experimental trials. They tended not to look at each other in the group situation except during intertrial break periods.

Summated Frequencies and Rate Data

In order to estmate the impact of the main independent variables upon the behaviors observed, raw frequencies were summed within behavioral categories across subjects and teams by experimental condition. For nonverbal behaviors, only data from persons 'C' and 'B' were summed. This was done to compensate for the fact that person 'A' sat behind a partition that blocked the recording camera's view during the individual viewing conditions. Summing across persons 'A', 'B', and 'C' in group condition would have artificially inflated the relative number of behaviors recorded for this condition.

Summary data are presented in tabular form within Attachment C. Verbal behavioral categories are presented first, followed by nonverbal categories. Within each category, higher confidence ratings are discussed first, followed by more tentative findings. Only simple main effects are described. No interaction effects or inferential statistics were calculated given the exploratory nature of the rating system and the low N observed. Results should, therefore, be viewed as impressionistic and speculative.

<u>Verbal Commands</u> (Table 1). Verbalizations that were overtly directed at coordinating activities of fellow team members (e.g., "Tom and Kathy, get that red triangle on row 8"; "Jim, get that blue rectangle and then meet us down on row 10") were most common. More verbal commands were recorded under conditions of low time stress (frequency = 661) than under high time

stress (frequency = 490). This difference could be accounted for in part by the greater amount of time available per trial when tasks were presented at a slow rate (60 tasks in 328 seconds) rather than at a fast rate (60 tasks in 164 seconds). Adjusting for this difference in total time available by converting raw frequencies into ratios of commands per time unit led to the finding that verbal command rates were actually higher in the high time stress condition (rate = 2.78 commands/minute) than the low time stress condition (rate = 1.88 commands/minute). Because most commands related to information being presented on the video screen, verbal communication appeared to be driven by task presentation rate. Verbal command rates were also slightly higher for teams whose members were isolated (rate = 2.3 commands/minute) than those whose members were clustered (rate = 2.0 commands/minute).

Positive and Negative Verbal Comments (Tables 2 and 3). These categories were originally designed to capture signs of group conflict and group solidarity. However, all teams observed were highly congenial with little or no interpersonal conflict observed. Observed affective verbalizations related primarily to comments about overall performance scores rather than interpersonal relations. Because the vast majority of comments occurred at the end of trials and during intertrial break periods, totals for these two time periods have been combined. Results suggest that team members were more willing to take the time to comment on positive outcomes under conditions of low time stress (frequency = 37) than for high time stress (frequency = 16) and made more negative comments when physically isolated from one another (frequency = 45) than under proximic conditions (frequency = 25).

Questions (Table 4). The rater reported particular difficulty with this category because many oberved questions contained implied commands (e.g., "Julie, want to do [row] 9?"). Questions were also of two general types--those referring to directions (e.g., "Do we want to go for the blue circle at the bottom?"), and those making reference to needed knowledge (e.g., "How many seconds do you have to have left?"). This category obviously needs to be refined. The only trend noted was more questions being asked at a higher rate of occurrence during the slow, low time stress condition

(frequency = 75, rate = .21 questions/minute) as compared to the fast, high time stress condition (frequency = 22, rate = .12 questions/minute).

Other Verbal Activity. Simple statements of fact (e.g., "Red triangle [coming] on [row] 9") were not tallied within the present rating system. This was probably a mistake since they accounted for close to half of the verbalizations observed. Automated sampling of total team verbalizations was used in the broader study of which this was a part. It would be of interest to compare total talking scores per condition. Unfortunately, these were not yet available at the time of the writing of this report.

Gesturing (Tables 5 and 6). The most common form of gesturing was pointing at the video display. This occurred almost exclusively within the group, large screen condition (frequency = 40, rate = .15 gestures/minute compared to frequency = 4, rate = .01 gestures/minute in the isolated condition). This was a clear case of an additional communication channel being available to team members sitting in a group rather than in isolated cubicles.

Self Manipulations (Tables 7 and 8). This category was originally included to detect signs of tension. While total frequencies are different between high and low time stress conditions, these differences disappear when adjustments are made for total available sampling time. Rates of responding remained fairly stable across experimental conditions.

Posture Changes (Table 9). This category included shifts in forward and backward lean as well as stretching. The latter subcategory was the most frequently occurring. It was observed almost exclusively at the end of trials during "break" periods and appeared more frequently in the group setting (frequency = 56) as compared to the isolated setting (frequency = 43). Yawning invariably was accompanied by stretching and was eventually dropped as a separate category.

Laughing/Smiling (Tables 10 and 11). Laughing and smiling behavior can be a sign of social affiliation or a sign of tension release. Frequency and rate of occurrence was higher in the high time stress condition (frequency = 60, rate = .34 occurrences/minute) than the low time stress

condition (frequency = 84, rate = .24 occurrences/minute). Laughing also occurred more frequently and at a slightly higher rate in the group (frequency = 85, rate = .32 occurrences/minute) than isolated (frequency = 59, rate = .22 occurrences/minute) settings. Like yawning and stretching, laughing may be affected by behavioral contagion in group settings.

Team Performance Scores

It was of interest to compare team performance measures with the verbal and nonverbal behaviors observed. Total points accumulated by teams during the trials observed (the first four trials of each experimental session) are shown within Attachment D. Total points accumulated were higher in the graphic (frequency = 6529) as opposed to the alphanumeric (frequency = 5562) display conditions and higher in the low time stress (frequency = 6585) as opposed to the high time stress (frequency = 5506) conditions. However, it is interesting to note that the rate of points accumulated followed a somewhat different pattern. While point accumulation rates remained higher for graphic (rate = 24.7 points/minute) than alphanumeric (rate = 21.1 points/minute) displays, team efficiency appeared to be higher during the fast, high time stress condition (rate = 31.2 points/minute) as opposed to the slow, low time stress condition (rate = 18.7 points/minute). The latter finding should be qualified by the realization that opportunities were limited under low time stress to perform optimally because of the slower presentation of tasks for processing.

In order to facilitate a comparison between team performance scores and the two behavioral categories that showed the most promise for tracking group decision making (i.e., verbal commands and gesturing), a limited set of Pearson product moment correlation coefficients was calculated. The number of verbal commands issued by emergent leaders (within the five teams showing a consistent leader) was correlated with total points accumulated on a trial by trial basis. These correlations were computed separately for high and low time stress conditions. Results indicated a modest positive correlation between these two variables under high time stress ($R = 15$, n.s.) and a negative correlation under low stress conditions ($R = -.33$, $p < .05$). This finding may be related to a recent finding reported by

Driskell (1986) that showed subordinates yielding more to leader influence under high stress conditions. However, without knowing what task choices were available to teams when the leader issued his/her commands, there is no way of knowing whether (s)he gave good advice or bad. Without knowing what individual responses were made, there is no way of knowing whether team members complied or resisted influence. Had a view of the video display been incorporated within the videotaped sessions, this kind of analysis could have been performed.

A correlation was also calculated between the number of gestures observed in teams working in the proximic seating arrangement and their total performance scores. Results indicated a modest negative relationship between these two variables ($R = -.23$, $p = .07$). It may be that those individuals pointing to the large screen display used this behavior as a substitute for clear verbal commands (there were slightly fewer commands issued in the group seating conditions compared to isolated group conditions). Given that team members remained focused upon their video displays, these gestures may have gone unnoticed by their intended targets.

Discussion

The primary motivation to conduct the present pilot was to (1) explore the use of new behavioral observation techniques to determine their utility within team decision making research, and (2) collect preliminary data that could be used to design future experiments dealing with communication issues within distributed decision making environments.

The rapid development of a behavioral rating technique and preliminary analyses of videotaped sessions proved to be a useful exercise for refining our thinking about observational techniques and narrowing in on a set of observable behaviors related to team decision making. For example, we now know that verbal command rates appear to be driven by information display rates. We also now know that leaders tend to emerge spontaneously in most groups to help coordinate task activity. Their effectiveness may be dependent upon stressors that force other team members to adhere to their directives. We have seen pointing behavior increase in response to gro p

seating and large screen displays, but have found this to be a counterproductive behavior.

While many of the behaviors sampled simply did not vary in a way that would help explain team performance outcomes, observing these behaviors across teams helped paint a picture of team operating climates that would have been missing had they not been recorded. For example, we know from the kind of positive and negative comments made that groups were generally cooperative and task oriented. We also know from attending to posture changes that team members were intently focused upon their video displays and did not move much during experimental trials. We also know from attending to posture changes that team members were intently focused upon their video displays and did not move much during experimental trials. We also discovered that intertrial SWAT rating periods were used by teams for a number of other activities including stretching and commenting on performance. Yawning, laughing, and stretching were more likely to be contagious during this period in proximic seating arrangements.

We have also learned something from our omissions. It is clear that we underestimated the richness of verbal behaviors in developing our rating scheme. More attention needs to be paid to the exchange of nonaffective, nondirective information. More attention also needs to be paid to the kinds of question team members ask each other.

It is believed that direct observation of group decision making activities is an important adjunct to the study of group performance outcomes. Observation of team communication patterns can provide insight into problem solving and decision making processes that are inaccessible within single person problem solving tasks. It is recommended that behavioral observation be continued and expanded in future team research. A detailed listing of recommendations for future research may be found in Attachment E.

References

Bales, R. F. (1950). Interaction process analysis: A method for the study of small groups. Cambridge, Massachusetts: Addison-Wesley.

Brown, C. E., and Leupp, D. G. (1985). Team performance with large and small screen displays. Armstrong Aerospace Medical Research Laboratory technical report number TR-85-033. (Available from NTIS.)

Driskell, J. E. (1986). Human performance in hostile environments. In Independent Research and Independent Exploratory Development Programs, Annual Report FY85 (pp. 13-28), NTSC TR-85-003. Orlando, Florida: Naval Training Systems Center.

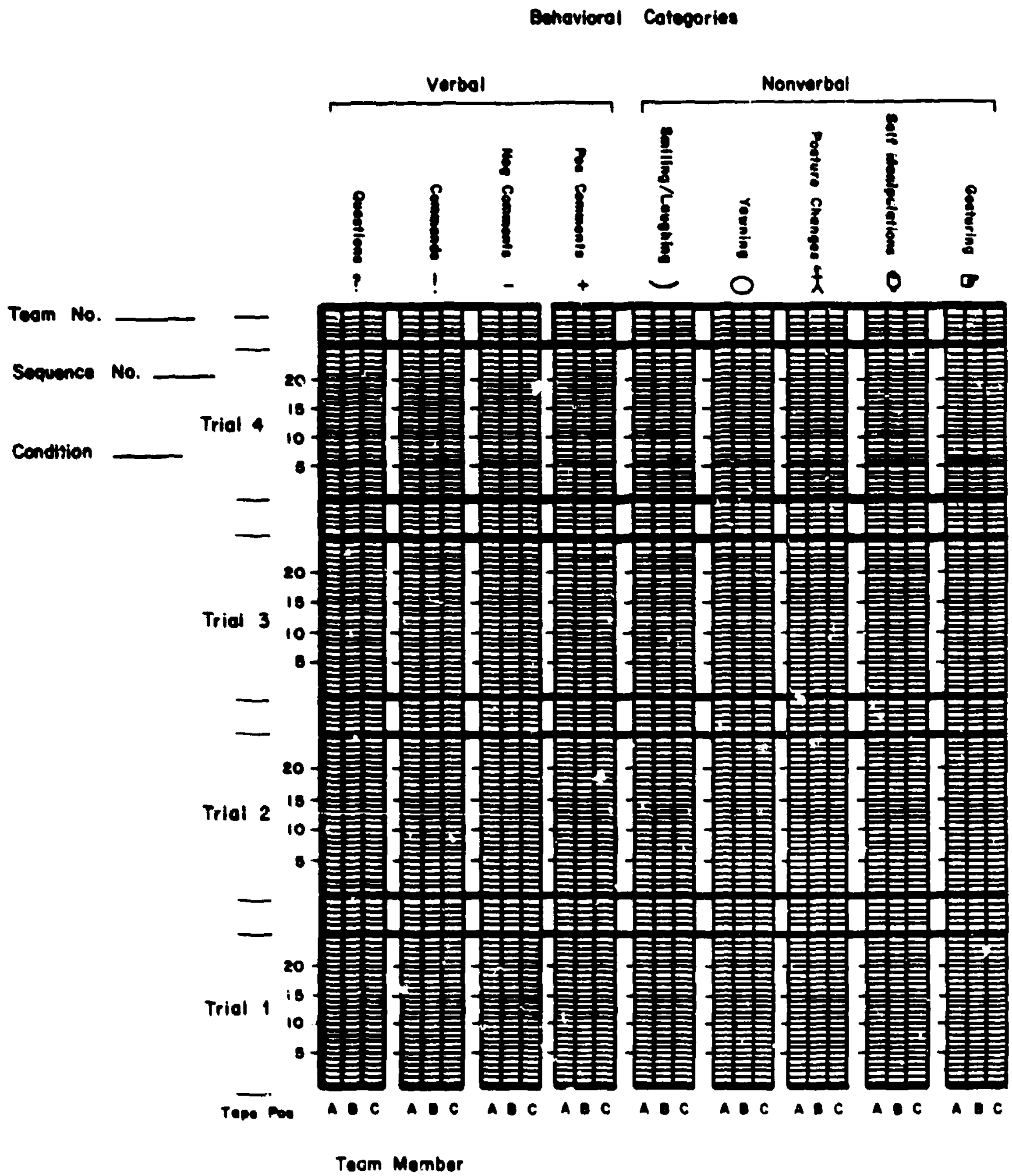
McNeese, M. D. and Brown, C. E. (1986). Large group displays and team performance: An evaluation and projection of guidelines, research, and technologies. Armstrong Aerospace Medical Research Laboratory technical report number TR-86-035.

Pattipati, K. R., Kleinman, D. L., and Ephrath, A. R. (1982). A dynamic decision model of human task performance. University of Connecticut Technical Report, TR-82-1, Department of EEEC.

Reid, G. B. (1982). Subjective workload assessment: A conjoint scaling approach. Paper presented at the annual meeting of the Aerospace Medical Association, Bal Harbour, Florida, May, 1982.

Attachment A

OBSERVATIONAL RATING SHEET



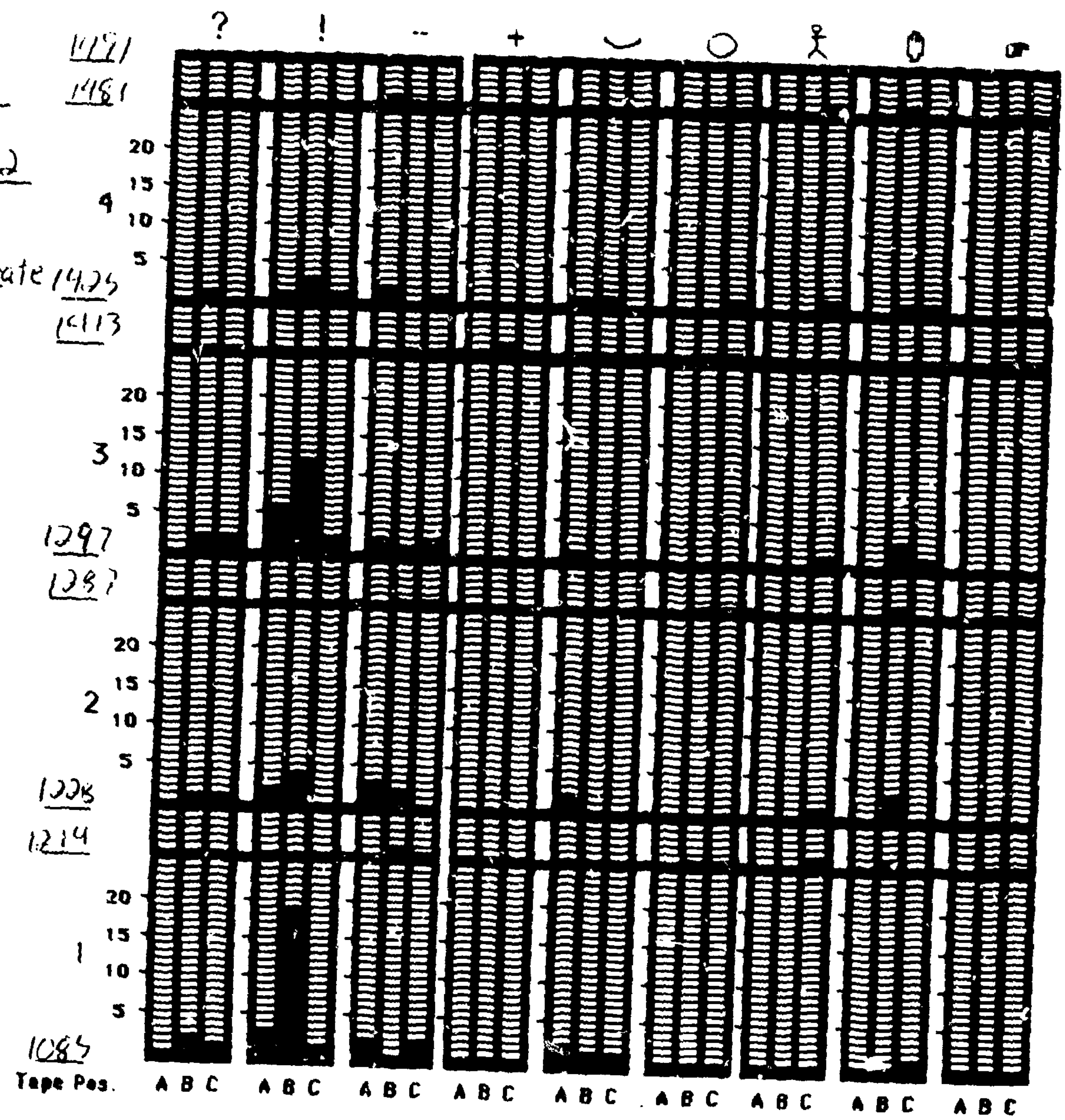
Attachment B

RAW DATA SHEET EXAMPLE

Team No. 10

Sequence No. 2

Condition separate



Attachment C

TABULARIZED FREQUENCY AND RATE DATA

TABLE 1. VERBAL COMMANDS*

(OCCURRENCES PER GROUP)

			Group Condition									
Team Number			10	11	12	13	14	15	17	18		
		Trial										
Alpha	Fast	4th	13	8	6	6	8	5	8	12	Total	121
		2nd	5	9	3	8	8	9	5	8		
	Slow	3rd	12	13	6	9	12	10	9	11	Total	142
		1st	6	10	4	4	10	10	8	8		
											Alpha Total	263
Graphics	Fast	4th	0	11	3	6	9	4	14	8	Total	109
		2nd	3	8	6	5	9	4	10	9		
	Slow	3rd	5	16	6	8	18	6	21	10	Total	164
		1st	6	14	6	8	12	4	11	13		
											Graph Total	273
											Group Total	536

			Separate Condition									
Team Number			10	11	12	13	14	15	17	18		
		Trial										
Alpha	Fast	4th	5	4	7	9	7	9	12	19	Total	116
		2nd	6	4	7	7	4	10	9	7		
	Slow	3rd	20	6	9	12	16	10	10	11	Total	176
		1st	23	0	6	8	11	15	10	9		
											Alpha Total	292
Graphics	Fast	4th	7	13	11	3	11	9	7	11	Total	144
		2nd	9	8	9	10	9	8	11	8		
	Slow	3rd	9	13	10	11	12	13	9	10	Total	179
		1st	9	5	16	13	16	13	12	8		
											Graph Total	323
											Separate Total	615

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	615	536	555	596	490	661
Frequency Per Minute	2.32	2.03	2.102	2.25	2.78	1.877

*Observations made on first four trials of each experimental session: data for group 15, group graphics condition, reflects an underestimation due to poor audio recording level for person 'A'.

TABLE 2. POSITIVE COMMENTS DURING AND BETWEEN TRIALS*
(OCCURRENCES PER GROUP)

Group Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	3	0	0	0	0	0	0	Total	4
		2nd	0	0	0	0	1	0	0	0		
	Slow	3rd	5	0	0	0	0	2	1	1	Total	10
		1st	0	0	0	1	0	0	0	0		
											Alpha Total	14
Graphics	Fast	4th	0	0	0	0	0	0	0	0	Total	2
		2nd	0	0	0	0	0	0	2	0		
	Slow	3rd	0	0	0	0	0	1	1	1	Total	7
		1st	0	1	0	0	2	1	0	0		
											Graph Total	9
											Group Total	23

Separate Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	1	0	0	1	0	0	0	Total	2
		2nd	0	0	0	0	0	0	0	0		
	Slow	3rd	1	0	0	1	0	1	1	0	Total	8
		1st	0	0	1	0	0	2	0	1		
											Alpha Total	10
Graphics	Fast	4th	0	1	0	1	0	2	0	0	Total	8
		2nd	0	2	0	0	0	1	1	0		
	Slow	3rd	3	1	1	0	0	0	0	0	Total	12
		1st	2	1	3	0	1	0	0	0		
											Graph Total	20
											Separate Total	30

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	30	23	24	29	16	37

*Observations made on first four trials of each experimental session: data for group 15, group graphics condition, reflects an underestimation due to poor audio recording level for person 'A'.

NEGATIVE COMMENTS DURING AND BETWEEN TRIALS*
(OCCURRENCES PER GROUP)

Group Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	0	1	0	0	0	0	1	Total	5
		2nd	2	0	0	0	0	0	0	1		
	Slow	3rd	3	2	0	0	1	0	0	0	Total	8
		1st	2	0	0	0	0	0	0	0		
											Alpha Total	13
Graphics	Fast	4th	0	0	0	0	0	0	1	0	Total	5
		2nd	1	0	0	0	0	3	0	0		
	Slow	3rd	0	0	0	0	1	0	2	0	Total	7
		1st	2	1	0	0	0	0	0	1		
											Graph Total	12
											Group Total	25

Separate Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	4	2	0	0	0	0	2	0	Total	15
		2nd	5	0	1	1	0	0	0	0		
	Slow	3rd	5	1	0	0	0	0	1	0	Total	14
		1st	5	0	0	0	0	1	1	0		
											Alpha Total	29
Graphics	Fast	4th	2	0	0	0	0	0	0	0	Total	6
		2nd	2	0	0	1	0	0	1	0		
	Slow	3rd	1	3	0	1	0	0	1	0	Total	10
		1st	2	0	1	1	0	0	0	0		
											Graph Total	16
											Separate Total	45

	Separate	Versus	Group	Alpha	Versus	Graph	Fast	Versus	Slow
Occurrences	45		25	42		28	31		39

*Observations made on first four trials of each experimental session: data for group 15, group graphics condition, reflects an underestimation due to poor audio recording level for person 'A'.

TABLE 4. QUESTIONS*
(OCCURRENCES PER GROUP)

Group Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	0	0	0	0	0	0	0	Total	3
		2nd	1	0	0	0	0	2	0	0		
	Slow	3rd	7	1	0	0	1	1	1	0	Total	22
		1st	3	0	0	1	2	1	4	0		
											Alpha Total	25
Graphics	Fast	4th	0	1	1	0	0	0	0	0	Total	11
		2nd	2	0	0	0	0	0	7	0		
	Slow	3rd	2	1	0	0	2	2	2	0	Total	20
		1st	2	3	0	0	2	1	3	0		
											Graph Total	31
											Group Total	56

Separate Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	1	0	0	0	1	0	0	0	Total	5
		2nd	2	0	0	0	1	0	0	0		
	Slow	3rd	4	2	1	0	2	0	0	0	Total	18
		1st	3	2	0	0	3	0	1	0		
											Alpha Total	23
Graphics	Fast	4th	0	0	0	0	1	0	0	0	Total	3
		2nd	1	0	0	1	0	0	0	0		
	Slow	3rd	0	0	0	0	3	0	2	0	Total	15
		1st	2	3	1	2	0	1	1	0		
											Graph Total	18
											Separate Total	41

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	41	56	48	49	22	75
Frequency Per Minute	0.15	0.21	0.181	0.185	0.12	0.213

*Observations made on first four trials of each experimental session: data for group 15, group graphics condition, reflects an underestimation due to poor audio recording level for person 'A'.

TABLE 5. GESTURING*

(OCCURRENCES PER GROUP)

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	0	0	0	1	0	0	0	Total	7
		2nd	4	1	0	0	0	1	0	0		
	Slow	3rd	1	0	1	0	1	2	0	0	Total	16
		1st	8	1	0	0	0	2	0	0		
											Alpha Total	23
Graphics	Fast	4th	0	0	1	0	1	0	0	3	Total	6
		2nd	0	0	0	0	0	0	0	1		
	Slow	3rd	0	0	0	0	0	0	0	5	Total	11
		1st	1	1	0	0	0	1	0	3		
											Graph Total	17
											Group Total	40

Separate Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	0	0	0	0	0	0	1	Total	1
		2nd	0	0	0	0	0	0	0	0		
	Slow	3rd	0	0	0	0	0	0	0	0	Total	0
		1st	0	0	0	0	0	0	0	0		
											Alpha Total	1
Graphics	Fast	4th	1	0	0	0	0	0	0	0	Total	2
		2nd	0	0	0	0	0	0	0	1		
	Slow	3rd	0	0	1	0	0	0	0	0	Total	1
		1st	0	0	0	0	0	0	0	0		
											Graph Total	3
											Separate Total	4

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	4	40	24	20	16	28
Frequency Per Minute	0.01	0.15	0.090	0.07	0.09	0.079

*Observations made on first four trials of each experimental session: data for based on obervation of persons 'B' and 'C' only; person 'A' not visible in separate condition.

TABLE 6. GESTURES BETWEEN TRIALS*
(OCCURRENCES PER GROUP)

Group Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
	Fast	4th	0	0	0	0	0	0	0	0	Total	0
Alpha		2nd	0	0	0	0	0	0	0	0		
	Slow	3rd	0	0	0	0	0	0	0	0	Total	7
		1st	3	0	0	3	1	0	0	0		
											Alpha Total	7
	Fast	4th	0	0	0	0	0	0	0	0	Total	0
Graphics		2nd	0	0	0	0	0	0	0	0		
	Slow	3rd	0	0	0	0	0	0	0	0	Total	3
		1st	0	0	0	0	1	0	0	2		
											Graph Total	3
											Group Total	10

Separate Condition

Team Number		Trial	10	11	12	13	14	15	17	18		
	Fast	4th	0	0	0	0	0	0	0	0	Total	0
Alpha		2nd	0	0	0	0	0	0	0	0		
	Slow	3rd	0	0	0	0	0	0	0	0	Total	0
		1st	0	0	0	0	0	0	0	0		
											Alpha Total	0
	Fast	4th	0	0	0	0	0	0	0	0	Total	2
Graphics		2nd	0	2	0	0	0	0	0	0		
	Slow	3rd	0	0	0	0	0	0	0	0	Total	0
		1st	0	0	0	0	0	0	0	0		
											Graph Total	2
											Separate Total	2

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	2	10	7	5	2	10

*Observations made on first four trial of each experimental session: data based on observation of persons 'B' and 'C'; person 'A' invisible during separate condition.

TABLE 7. SELF MANIPULATION*
(OCCURRENCES PER GROUP)

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	1	0	2	1	3	1	0	0	Total	26
		2nd	1	0	1	3	5	1	1	6		
	Slow	3rd	2	5	5	4	2	6	0	7	Total	53
		1st	4	3	5	3	3	2	1	1		
											Alpha Total	79
Graphics	Fast	4th	2	0	3	1	3	2	1	6	Total	31
		2nd	2	3	2	2	0	1	0	3		
	Slow	3rd	5	4	8	7	3	1	2	7	Total	68
		1st	3	3	1	7	4	3	2	8		
											Graph Total	99
											Group Total	178

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	2	3	0	2	0	1	3	3	Total	32
		2nd	3	6	2	1	0	1	4	1		
	Slow	3rd	4	4	2	3	4	6	3	3	Total	43
		1st	1	3	4	2	0	1	1	2		
											Alpha Total	75
Graphics	Fast	4th	6	2	1	2	0	0	0	2	Total	23
		2nd	4	2	1	0	1	0	0	2		
	Slow	3rd	3	3	8	3	2	1	0	5	Total	45
		1st	2	0	6	2	2	0	5	3		
											Graph Total	68
											Separate Total	143

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	143	178	154	167	112	209
Frequency Per Minute	0.54	0.67	0.583	0.63	0.63	0.593

*Observations made on first four trials of each experimental session: data based on observations of persons 'B' and 'C' only; person 'A' not visible in separate condition.

TABLE 8. SELF MANIPULATION BETWEEN TRIALS*
(OCCURRENCES PER GROUP)

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	1	0	0	0	0	1	2	1	Total	12
		2nd	1	1	0	2	2	0	1	0		
	Slow	3rd	1	5	0	0	0	1	1	2	Total	17
		1st	3	1	0	0	1	0	2	0		
											Alpha Total	29
Graphics	Fast	4th	2	1	2	0	0	1	1	0	Total	21
		2nd	3	2	3	1	2	1	0	2		
	Slow	3rd	2	1	1	0	0	1	0	0	Total	17
		1st	2	3	1	0	2	0	2	2		
											Graph Total	38
											Group Total	67

Team Number		Trial	10	11	12	13	14	15	17	18		
Alpha	Fast	4th	1	0	0	0	0	0	4	0	Total	13
		2nd	1	5	1	0	0	0	1	0		
	Slow	3rd	0	2	1	0	1	1	2	1	Total	15
		1st	1	1	0	2	1	0	2	0		
											Alpha Total	28
Graphics	Fast	4th	2	2	1	2	0	1	1	0	Total	16
		2nd	1	1	2	1	1	0	0	1		
	Slow	3rd	3	3	1	3	1	0	0	1	Total	18
		1st	1	1	0	0	2	0	2	0		
											Graph Total	34
											Separate Total	62

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	62	67	57	72	62	67

*Observations made on first four trials of each experimental session: data based on observation of persons 'B' and 'C' only; person 'A' not visible during separate condtion.

TABLE 9. POSTURE CHANGES BETWEEN TRIALS*
(OCCURRENCES PER GROUP)

Team Number		Trial	10	11	12	13	14	15	17	18		
	Fast	4th	1	1	1	0	1	0	0	1	Total	14
Alpha		2nd	2	3	2	0	1	0	0	1		
	Slow	3rd	0	3	1	0	0	1	1	1	Total	13
		1st	1	3	0	0	1	0	1	0		
										Alpha	Total	27
	Fast	4th	4	2	2	0	1	0	0	0	Total	16
Graphics		2nd	2	2	1	0	1	0	1	0		
	Slow	3rd	1	3	1	0	0	0	0	0	Total	13
		1st	1	2	1	0	1	1	2	0		
										Graph	Total	29
										Group	Total	56

Team Number		Trial	10	11	12	13	14	15	17	18		
	Fast	4th	1	0	0	0	2	0	0	0	Total	8
Alpha		2nd	0	4	0	0	0	0	1	0		
	Slow	3rd	0	2	2	0	2	0	1	0	Total	11
		1st	1	0	0	0	2	0	1	0		
										Alpha	Total	19
	Fast	4th	3	2	3	0	0	1	0	0	Total	14
Graphics		2nd	1	2	2	0	0	0	0	0		
	Slow	3rd	1	2	1	0	1	0	0	0	Total	10
		1st	2	2	0	0	1	0	0	0		
										Graph	Total	24
										Separate	Total	43

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	43	56	46	53	52	47

*Observations made on first four trials of each experimental session: data based on observation of persons 'B' and 'C' only; person 'A' not visible in separate condition.

TABLE 10. LAUGHTER*
(OCCURRENCES PER GROUP)

Team Number		Trial	Group Condition 10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	2	0	0	0	0	2	4	Total	14
		2nd	1	2	0	0	0	0	1	2		
	Slow	3rd	0	9	0	2	0	0	2	5	Total	29
		1st	1	4	0	0	0	0	2	4		
											Alpha Total	43
Graphics	Fast	4th	0	9	0	0	0	0	2	1	Total	16
		2nd	2	0	0	0	0	0	2	0		
	Slow	3rd	2	4	0	0	0	0	4	2	Total	26
		1st	2	2	0	2	2	0	4	2		
											Graph Total	42
											Group Total	85

Team Number		Trial	Separate Condition 10	11	12	13	14	15	17	18		
Alpha	Fast	4th	2	0	0	1	0	0	0	0	Total	5
		2nd	2	0	0	0	0	0	0	0		
	Slow	3rd	1	0	0	0	0	0	2	0	Total	9
		1st	4	2	0	0	0	0	0	0		
											Alpha Total	14
Graphics	Fast	4th	11	3	0	2	0	0	1	0	Total	25
		2nd	5	1	0	2	0	0	0	0		
	Slow	3rd	8	0	0	1	0	0	0	1	Total	20
		1st	1	5	0	1	0	0	2	1		
											Graph Total	45
											Separate Total	59

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	59	85	57	87	60	84
Frequency Per Minute	0.22	0.32	0.215	0.32	0.34	0.238

*Observations made on first four trials of each experimental session: data for group 15, group graphics condition, reflects an underestimation due to poor audio recording level for person 'A'.

TABLE 11. LAUGHTER BETWEEN TRIALS*
(OCCURRENCES PER GROUP)

Team Number		Trial	Group Condition 10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	4	0	0	0	0	0	0	Total	9
		2nd	0	0	0	0	3	0	2	0		
	Slow	3rd	0	3	0	1	0	0	0	0	Total	7
		1st	2	0	0	0	0	0	1	0		
											Alpha Total	16
Graphics	Fast	4th	0	2	0	0	2	0	1	2	Total	15
		2nd	0	3	0	0	0	2	1	2		
	Slow	3rd	0	3	2	0	0	0	0	0	Total	10
		1st	0	2	0	0	1	0	2	0		
											Graph Total	25
											Group Total	41

Team Number		Trial	Separate Condition 10	11	12	13	14	15	17	18		
Alpha	Fast	4th	0	0	0	0	0	0	0	0	Total	7
		2nd	0	7	0	0	0	0	0	0		
	Slow	3rd	0	0	0	0	0	0	0	0	Total	1
		1st	0	0	0	0	1	0	0	0		
											Alpha Total	8
Graphics	Fast	4th	0	0	0	0	0	0	0	0	Total	2
		2nd	0	0	0	0	0	0	0	2		
	Slow	3rd	0	0	0	0	0	0	0	0	Total	2
		1st	0	1	0	0	1	0	0	0		
											Graph Total	4
											Separate Total	12

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Occurrences	12	41	24	29	33	20

*Observations made on first four trials of each experimental session: data for group 15, group graphics condition, reflects an underestimation due to poor audio recording level for person 'A'.

Attachment D

TABULARIZED PERFORMANCE MEASURES

GROUP PERFORMANCE*
(POINTS ACCUMULATED PER GROUP)
Group Condition

Team Number		Trial	10	11	12	13	14	15	17	18	
	Fast	4th	92	108	71	88	73	84	87	79	Total 1272
Alpha		2nd	51	74	76	89	74	72	87	67	
	Slow	3rd	87	85	115	99	102	106	98	124	Total 1520
		1st	62	102	90	105	95	49	97	104	
											Alpha Total 2792
	Fast	4th	99	87	91	105	105	87	61	98	Total 1510
Graphics		2nd	113	106	105	120	100	74	62	97	
	Slow	3rd	131	115	115	131	101	115	89	112	Total 1758
		1st	113	102	96	122	108	108	82	108	
											Graph Total 3268
											Group Total 6060

Separate Condition

Team Number		Trial	10	11	12	13	14	15	17	18	
	Fast	4th	61	51	99	87	39	83	97	79	Total 1199
Alpha		2nd	64	56	91	70	50	90	103	79	
	Slow	3rd	88	98	115	121	72	122	104	108	Total 1571
		1st	67	93	110	98	68	96	89	122	
											Alpha Total 2770
	Fast	4th	94	100	92	96	102	95	95	98	Total 1525
Graphics		2nd	88	85	83	80	101	96	109	111	
	Slow	3rd	122	114	115	77	116	121	98	124	Total 1736
		1st	100	97	99	108	113	107	98	126	
											Graph Total 3261
											Separate Total 6031

	Separate	Versus Group	Alpha	Versus Graph	Fast	Versus Slow
Points	6031	6060	5562	6529	5506	6585
Average	94.2	94.6	86.90	102.	86.0	102.8
Points Per Minute	22.8	22.9	21.06	24.7	31.2	18.70

*Data from first four trials of each experimental session only.

Attachment E

RECOMMENDATIONS FOR FUTURE RESEARCH

Throughout Appendix D, we have mentioned a number of methodological and procedural shortcomings that could easily be corrected in future studies. Suggested "fixes" are summarized below together with additional comments regarding future research directions.

Technological Fixes

1. Upgrade videotaping facility to allow multicamera, unobstructed views of team members in group and isolated conditions.

2. Incorporate views of team members' CRT displays within videotapes to provide information regarding their current focus of attention.

3. Computerize rating procedure and videotape handling procedures.

4. Maintain log of individual as well as group point accumulation within TRAP task.

Methodological Fixes

1. Unconfound large versus small screen displays from group versus isolated viewing.

2. Develop a more comprehensive behavioral observation system that includes a more detailed taxonomy of verbal behaviors.

3. Use multiple raters and extend training time to increase reliability of behavioral measures.

4. Supplement SWAT ratings with alternative measures of workload (i.e., physiological measures or embedded tasks).

5. Develop new measures of team performance efficiency.

6. Control for sex composition and prior acquaintance of team members.

7. Increase the number of teams observed and expand subject population.

8. Decrease dependence on within subject designs.

9. Increase breadth of gaming situations.

10. Consider field and survey research to supplement laboratory experiments.

Theoretical Concerns

1. Explore situational stressors in addition to time stress (e.g., information overload, conflict or ambiguity, situation severity, etc.).

2. Include a consideration of leadership variables (e.g., seating position, personality variables) in future team experiments.

3. Wherever possible, develop competing theoretical frameworks to allow critical tests among alternative hypotheses.

U.S. GOVERNMENT PRINTING OFFICE 1986